AD 747997

UNCLASSIFIED

Security Classification

DOCUMENT CONTROL DATA - R & D

(Security classification of title, body of abstract and indexing annotation must be entered when the overall report is classified)

1. ORIGINATING ACTIVITY (Corporate author) Applied Psychological Services, Inc. Science Center Wayne, Pennsylvania 19087	2a. REPORT SECURITY CLASSIFICATION UNCLASSIFIED 2b. GROUP

3. REPORT TITLE

Nonverbal and Culture Fair Performance Prediction Procedures. I. Background, Test Development, and Initial Results

4. DESCRIPTIVE NOTES (Type of report and inclusive dates)

Annual Report, July 1971-June 1972

5. AUTHOR(S) (First name, middle initial, last name)

Arthur I. Siegel
Brian A. Bergman

6. REPORT DATE	7a. TOTAL NO. OF PAGES	7b. NO. OF REFS
June 1972	77 + ix	67

8a. CONTRACT OR GRANT NO. N00014-71-C-0319 b. PROJECT NO. NR 153-334 c. d.	9a. ORIGINATOR'S REPORT NUMBER(S) 9b. OTHER REPORT NO.(S) (Any other numbers that may be assigned this report)

10. DISTRIBUTION STATEMENT

This report has been approved for public release and sale; its distribution is unlimited.

11. SUPPLEMENTARY NOTES	12. SPONSORING MILITARY ACTIVITY
	Personnel and Training Programs Office of Naval Research

13. ABSTRACT

The logic and initial results are described of a program into the development of unique measures for assessing the potential of "low aptitude" personnel for certain Navy rates. The logic is based on the conjecture that recruits who can learn a sample of the job requisites in a mini on-the-job training situation will demonstrate the same ability on the job. This is held to apply regardless of the recruit's low score on the usual classification tests. The initial and criterion tests are described and the correlations among the mini job learning test results and the usual Navy predictors are given. The results of a factor analysis of a questionnaire related to cultural deprivation are given, and the relationship of the derived cultural deprivation scores both to the usual Navy classification tests and the job learning tests is given.

ia

KEY WORDS	LINK A ROLE	LINK A WT	LINK B ROLE	LINK B WT	LINK C ROLE	LINK C WT
Assessment						
Tests and Measurement						
Culture Fair Tests						
Personnel Prediction						
Cultural Deprivation						
Differential Validity						
Performance Criteria						

ib

NONVERBAL AND CULTURE FAIR PERFORMANCE PREDICTION PROCEDURES

I. Background, Test Development, and Initial Results

Arthur I. Siegel
Brian A. Bergman

prepared for

Personnel and Training Programs
Psychological Sciences Division
Office of Naval Research

by

Applied Psychological Services, Inc.
Science Center
Wayne, Pennsylvania

under

Contract N00014-71-C-0319
NR 153-334

1C

June 1972

ABSTRACT

The logic and initial results are described of a program into the development of unique measures for assessing the potential of "low aptitude" personnel for certain Navy rates. The logic is based on the conjecture that recruits who can learn a sample of the job requisites in a mini on-the-job training situation will demonstrate the same ability on the job. This is held to apply regardless of the recruit's low score on the usual classification tests. The initial and criterion tests are described and the correlations among the mini job learning test results and the usual Navy predictors are given. The results of a factor analysis of a questionnaire related to cultural deprivation are given, and the relationship of the derived cultural deprivation scores both to the usual Navy classification tests and the job learning tests is given.

FOREWORD

In order to reach the goals of the present study, we depended on a number of people for guidance, assistance, and support. Chief Nathaniel Hamilton, of the North Chicago High School, served as an instructor/test administrator during the instructional and test administrative phase of the work. Mr. Lester Harmon, Superintendent of North Chicago Schools, kindly granted Chief Hamilton the permission to work with us.

Dr. Victor Fields, Office of Naval Research, provided field liaison and guidance during the initial phases of this research effort. Dr. Marshall Farr has continued to encourage our efforts and has provided helpful suggestions. LT Martin Boone provided assitance in administrative matters.

Chief J. W. Helt, of the Recruit Training Command's Classification Section at Great Lakes, handled subject assignments and administrative details. Chief R. G. Reynolds, also of the Recruit Training Command, provided technical assistance and arranged for Chief Hamilton to join our research team. LT C. H. Regner, of the Recruit Training Command, acted as our Liaison Officer. CAPT Symans, Commanding Officer, Recruit Training Command, gave our project support and encouragement.

LCDR D. S. Betz, Commanding Officer, Propulsion Engineering School, Great Lakes, permitted us to use his school's facilities. CWO R. E. Howery, Chief J. L. DeLong, and Chief H. E. Wilton, of the Machinist Mate A School at Great Lakes, all provided much needed technical assistance during the course of this project.

CWO J. T. Dubose, CWO D. Romanek, CWO J. McCool, and Chief W. C. Villo, of the Damage Control School at the Philadelphia Navy Yard, helped us to contruct the criterion tests to be used in the evaluation of the Fleet performance of our subjects. CDR May, Director of the Damage Control School, arranged for these members of his staff to consult with us.

Arthur I. Siegel
Brian A. Bergman

APPLIED PSYCHOLOGICAL SERVICES, INC.
June 1972

Preceding page blank

TABLE OF CONTENTS

Preceding page blank

Page

LIST OF TABLES

Table		Page

TABLE OF FIGURES

CHAPTER I

INTRODUCTION AND PURPOSE

Within the recent past, traditional verbal measures of assessment have come under both legal and psychometric scrutiny. The legal criticisms are interwoven with concepts related to discrimination and fair employment, while the psychometric criticisms have been largely, but not exclusively, involved with differential validity and "culture fairness." The present program is based on a concept related to "culture fairness" or "culture freeness" but may be more properly associated with a "culture loaded" description. By culture loaded, we mean performance prediction on the basis of instruments that have been loaded in the job culture of interest. It is contended that if a person can learn to perform a job sample, he can also learn to perform the total job. Accordingly, the demonstrated ability to learn selected job aspects is employed as a predictor of ability to learn to perform the total job. The job sample learning situations (mini tests) involve no written learning materials. Hence, they maximally simulate the on-the-job training situation in which a foreman instructs a journeyman in job performance.

Legal Aspects of Employment Discrimination

Chief Justice Warren Burger (1970), in writing the majority decision concerning the Duke Power vs. Griggs case, indicated that employment policies, even if nondiscriminatory in intent, which "... 'freeze' the status quo of prior discriminatory employment practices" (p. 5) cannot be maintained. He qualified this statement by indicating that a person need not be hired purely because he was once discriminated against, or because he is a minority group member. Congress just wishes "... the removal of artificial, arbitrary, and unnecessary barriers when the barriers" (p. 6) ... promote discrimination against certain groups of persons. If an employment practice is unrelated to job success, it is illegal. Just because a selection test was professionally developed does not mean that it can be used for employment purposes. Congress has said "... that any tests used must measure the person for the job and not the person in the abstract" (p. 12).

The above statement means that any employment tests used must be shown to be related to specific aspects of behaviors common to the job in question. If employers use tests or other instruments that measure abilities not required of the job, they are breaking the law. Ruch (1971) maintained that most intelligence tests measure nonrelevant aspects of jobs and are therefore illegal. He recommended the use of unique factor tests measuring specific job skills. Enneis (1969b) has arrived at a similar conclusion. Specifically, he favored relating selection standards to job requirements.

The American Psychological Association's Task Force on Employment Testing of Minority Groups (1969) listed most of the common sources of discriminatory bias of which employers should be aware:

1. inadequate recruiting in minority schools or neighborhoods

2. unfair preliminary selection by the receptionist

3. application blanks which are clinically interpreted

4. weighted application blanks which may not be valid for the minority group

5. interviews which are open to "... conscious and unconscious perceptual bias" (p. 640). Most interviewers look for subsequent interview facts which support their first impression, and they neglect facts which go against their first impressions.

6. promotion practices in which the supervisor's appraisal of the minority employee may be influenced by racial characteristics and not job performance

7. test anxiety, including unfamiliarity of the minority applicant (or employee) with testing or inadequate adherence to optimal testing conditions

8. test content (especially verbal content) which is mainly derived from middle class culture

9. unfair interpretation of test scores

10. test content which is unrelated to the job requirements.

The task force did not recommend the removal of cultural differences (e.g., culture free tests, analysis of covariance), but rather the identification from among the low scorers those who can overcome their lack of experience and who can profit from extra training. The responsibility for the extra training rests with the employer.

Problems in Establishing Transethnic Group Test Fairness

Krug (1966) listed and criticized some of the methods used to eliminate or control for test score differences so that minority group members seeking employment will not be discriminated against. The most primitive method for promoting equal employment opportunity is through the use of double standards (when the mean test score of the minority group is significantly lower than the mean test score of the majority group). In this instance, the employer uses a lower cutoff score for the minority applicants and avoids the exclusion of an "unfair" proportion of the minority group members from employment. Variations on this theme have been proposed by Thorndike (1971) and Darlington (1971). Thorndike maintained that ... "If one acknowledges that differences in average test performance... exist between populations A and B, then a judgment on test-fairness must rest on the inferences that are made from the test rather than on a comparison of mean scores in the two populations" (p. 63). Thorndike advocated setting two different cutting scores for different groups in order to achieve optimal fairness. Darlington (1971), on the other hand, would add a certain number of points to the lower (minority) group and then apply the same cutting score. The double standard method is, of course, no substitute for differential prediction of minority and nonminority performance.

"Culture free," "culture fair," and "culture equivalent" tests have also been proposed as methods for performing fair across groups assessment. The main assumption of proponents of "culture free" tests is that the test content has the same meaning for all cultures. Unfortunately, this assumption can seldomly be met. In addition, the content of such tests makes them irrelevant for application assessment purposes.

Another kind of test with more modest claims attached to it is the "culture fair" test. The "culture fair" test is assumed to contain a set of stimuli which are appropriate for at least two cultures. Krug suggests that "culture common" (p. 33) would be a better description of this kind of test than "culture fair."

"Culture equivalent tests" are different from "culture fair" tests because they are not predicated on common material, but rather on material which tests the same concept using stimuli appropriate to different cultures.

Boehm (1971) recently examined 13 research studies which reported either differential validity or single-group validity. Differential validity exists when, for two or more groups, significant differences exist between predictor-criterion correlations and for one or both groups these coefficients are significantly different from zero. In single-group validity, the difference between validity coefficients for the two groups is not statistically significant, and only one coefficient is significantly different from zero. Boehm's analysis indicated that:

1. Only 60 of 160 (37.5%) of the correlation coefficients reported in the 13 studies reviewed were statistically significant for either the white or the black ethnic group.

2. In five of 13 studies, the predictors were used as selection measures with a resultant restriction in range. These predictors will usually have low validity, and using them excludes a high proportion of Negroes at the beginning and reduces the number of subjects for this group.

3. Only seven of the correlations were associated with differential validity. Many of these plus the single group validities can be attributed to inadequate sample sizes. When the N of both Negroes and whites was above 100, there was no differential or single-group validity.

4. Job knowledge and performance test criteria tended to yield higher validity coefficients than supervisory ratings.

5. In 22 of 27 instances of overall validity (correlation coefficients of both groups significant), supervisory ratings or rankings were not used as performance criteria, but in 19 of 33 single-group validities a rating was used; therefore, single-group and both-group validity are associated with the type of criteria used.

Boehm concluded that single-group validity is associated with small sample size and a supervisory rating criterion. Both-group validity, though, is dependent upon large sample size and the use of performance of job knowledge tests as criteria. She therefore recommended that employers use more objective measures of employee performance rather than supervisory ratings or rankings. Bennett (1969) has arrived at a similar conclusion.

One study illustrating some of the problems inherent in using rating scales as performance criteria was performed by Flaugher, Campbell, and Pike (1969). Supervisory ratings were examined to determine if the ethnic group membership of the ratee and the rater influenced assessment. One-hundred sixty-eight Negro and 296 white medical technicians employed in VA hospitals were all rated at least twice and given a separate job knowledge test. On the nine rating scales considered collectively, whites were rated higher ($p < .01$). Whites also scored higher on the job knowledge test ($p < .01$). Generally, Negro supervisors rated Negro incumbents higher than did white supervisors (one-half a standard deviation). Negro supervisors, though, did not rate white incumbents higher or lower than white supervisors.

Bartlett and O'Leary (1969) presented four cases illustrating how validity coefficients can yield erroneous predictions when a single coefficient is used to predict performance for two ethnic groups.

The first situation is that in which there are significant differences between test scores and criterion scores for both groups. In this situation, the overall validity coefficient can be increased if the differences are in the same direction for both groups.

A second instance, also discussed by Guion (1966), Kirkpatrick, Ewen, Barrett, and Katzell (1967), and Linn and Werts (1971), is that in which the test scores for the caucasian group are higher, but the criterion scores are equal so that the performance of caucasians is overpredicted and the performance of Negroes is underpredicted.

The third case exists when there is a difference between the groups in performance, but no difference in the predictor scores. In this situation, the performance of the high performance group will be underpredicted and the performance of the low performance group will be overpredicted.

Finally, when there is a difference in opposing directions between the predictor and the criterion in both groups, the combined validity coefficient may be negative, yet differential prediction is positive for each group separately. Using an overall validity coefficient, in this situation, would result in the selection of the worst performers from each group.

Bartlett and O'Leary also discussed four instances in which differential validity can occur. Some of these overlap, to some extent, with their single coefficient cases. First, a test can be valid for one group and not for the other group, but the mean scores on the criterion and the predictor may be the same for both groups. If the combined validity coefficient were used (if it was significant), one might select better persons from the valid group and erroneously conclude that the performance of the nonvalid group was inferior. The solution to this problem is to seek other valid predictors for the nonvalid group. Another case is that in which the groups differ in average performance, but exhibit no significant difference in their predictor score differences. The third case is that in which differences exist between the means for the predictors, but no differences exist in the criterion means, resulting in differential validity for the two groups. A final and important case is that in which both the predictor and the criterion means are higher for one group. In this case, combining groups would increase the validity, but the increased validity is due to racial (or sex, etc.) differences, and use of this is illegal.

Bartlett and O'Leary concluded that, in addition to searching for valid predictors of minority group performance, employers should initiate training procedures to enhance the minority group's chance of success.

Einhorn and Bass (1971) demonstrated that prediction is not necessarily better in a group with a higher validity coefficient over a group with a lower validity coefficient. This situation obtains when the high validity group has a larger standard error of estimate than the low validity group.

Enneis (1969a) stressed the control of skewness to increase the fairness of prediction. When the variance of the scores of the minority group is less than the variance of the scores of the majority group and a favorable selection ratio exists (1 in 10), then more nonminorities will be hired. Also, if the score distribution of nonminorities is negatively skewed and the score distribution of minorities positively skewed, then the majority group will be favored.

Bennett (1969) and Enneis (1969c) both indicated that concurrent validity studies do not give a conservative estimate of predictive validity. On the other hand, many psychologists feel that, because of the restricted score range in concurrent validity studies, the correlation estimate is an underestimate of the predictive validity coefficient.

Studies into Differential Validity

Foley (1971) wished to investigate whether or not the Officer Qualification Test (OQT) used by the Navy was biased against Negroes. Foley posited that the OQT is unfair if the regression equation used to predict success for Caucasians underpredicts the performance of Negroes. The Negro sample was compared with a matched white sample and with an unselected white sample. The OQT predicted performance in school (grade point average) for Negroes ($r = .29$) and for a matched Caucasion group ($r = .48$). Foley demonstrated that use of the Caucasion regression equation for Negroes resulted in a slight overprediction of Negro performance in Officer Candidate School.

Plag and Goffman (1967) found that educational level was a better predictor of performance in the Air Force (four year effectiveness, semiannual marks, advancement, disciplinary and commendatory actions, and adjustment ratings) than the Armed Forces Qualification Test (AFQT). Race was found to be an ineffective predictor of performance.

Baehr, Saunders, Froemel, and Furcon (1971) conducted a large scale differential validity study of policemen in the city of Chicago. Although there are several serious statistical errors in this study, some of the findings are worth noting. These investigators found that the best overall prediction was obtained for the black group, and that different tests (some overlapping) predicted performance across both racial groups.

Lopez (1966) suggested that different standards be used for prediction of performance across subcultural groups. He found that Negro toll collectors scored lower on several tests, but performed equal to whites on the job. Moreover, different predictors were related to performance across racial groups.

Mitchell and Albright (1968) conducted a large scale validational study in a large southern plant using the Wonderlic Intelligence Test and a Biographical Information Blank as predictors. The criteria used by Mitchell and Albright were supervisory ratings, rankings, and turnover. These investigators found that the Wonderlic failed twice as many blacks (54 per cent) as whites (27 per cent) and that it was not a valid predictor for either racial group. Scores on the Biographical Information Blank, though, correlated .30 with turnover.

In a study performed at IBM (Wollowick, Greenwood, & McNamara, 1969), 576 administrative personnel were given "in house" tests of vocabulary, nonverbal reasoning, and arithmetic. Salary and rankings were used as performance criteria. Black employees (N = 60) were matched with three separate samples of white employees in three different ways: (a) job and demographic variables, (b) supervisory evaluations, and (c) adjusted salary. For the second (b) and third (c) matched samples, whites tended to have significantly higher test scores than blacks, even though they were matched on performance and salary. Also, the white groups yielded higher validity coefficients than the black group.

Grant and Bray (1970), in a recent study, used a task proficiency after training criterion because they considered task proficiency to be uninfluenced by supervisory bias, peer pressures, or motivation. These writers asked five telephone companies to hire 100 employees, one-half of whom were black, and one-half of whom did not meet recommended standards on a battery of written tests. Each employee was sent through a training program consisting of seven levels. At each training level, an employee took a pre-test. If he passed the pre-test, he went on to the next level of training. If he failed the pre-test, he took training at that level, followed by a post-test. The highest training level passed was used as the best performance criterion. All the prediction instruments correlated significantly (mostly at $p < .01$) with the highest level passed in the training program. The minority and nonminority correlations were all virtually identical. The SCAT and the Abstract Reasoning Test gave a multiple of R of .49, and these were subsequently used for hiring purposes.

Ruda and Albright (1968) conducted a study which illustrates some of the problems inherent in using a single validity coefficient to predict across racial groups. These investigators found that the Wonderlic was weighted more heavily (first hurdle) in the prediction of turnover than a BIB (second hurdle). The Wonderlic correlated -.34 with turnover for whites and +.10 for blacks. This situation results in the employment of the whites who are most likely to turnover and essentially chance prediction for blacks. The weighted application blank score correlated much better with turnover than the Wonderlic, as evidenced by a .24 correlation for blacks and a .18 correlation for whites.

Qualitative Differences in Intellectual Functioning and Performance

Rimland (1969) using an idea similar to Jensen's (1969) differentiated between abstract intelligence ("g") (the ability to manipulate symbols and events mentally) and practical intelligence (the ability to sustain or perform simple tasks which simulate a job). Rimland posited that these intellectual types are inversely related to one another so that an individual who is high on one will not be high on the other. Most traditional tests of an abstract nature represent acceptable predictors of academic success, while practical performance tests are better predictors of job performance. Rimland found that practical performance tests correlated .19 to .37 with job performance in the Navy, while the highest correlation between the AFQT and job performance was .22.

McFann (1969) indicated, after reviewing previous research, that the differences between high and low aptitude men in Basic Combat Training was not as marked on motor skills and proficiency tests. On these latter tests, category four personnel usually met standard. In a project SPECTRUM study, men representing high, middle, and low aptitude groups were selected and individualized training instituted using videotape, a one-to-one student-teacher ratio, and specialized training. In some tasks, low aptitude men reached standard, but took from twice to four times as long, and in other cases, they failed to master the material at all. McFann also found the high aptitude group to learn equally well with lecture or individualized training, while low aptitude groups learned well with individualized training, but not with lecture; therefore, aptitude interacts with method of instruction.

Taylor, Montague, and Hauke (1970) were critical of the Army's lock-step training procedures. They indicated that this type of training makes it difficult to train high and low aptitude personnel together. The high aptitude students are held back because the training is not enough of a challenge, while the low aptitude students fail to learn because the material is too difficult. These researchers developed a miniaturized training sequence utilizing a variety of different procedures. The subjects used were 350 low aptitude recruits, 190 middle aptitude recruits, and 180 high aptitude recruits. The best training approach for high aptitude recruits was one without structure. High aptitude recruits should be given the objectives of training, allowed to choose their own study methods, and make their own decision as to when they are ready for testing. Middle aptitude subjects derive

the most benefit from the same kind of training given to high aptitude subjects except that they prefer the presence of a live monitor. The training method which is most beneficial to low aptitude subjects has:

1. complete structure
2. instruction presented in small steps
3. a slow rate of presentation
4. a high rate of repetition
5. an elementary language level
6. content presented in a functional context with provision for practice
7. a live instructor
8. constant prompting and feedback
9. meaningful extrinsic motivators

Moore, MacNaughton, and Osburn (1969) indicated that nonverbal tests are not necessarily the least biased against minority groups. These writers gave both verbal and spatial (nonverbal) tests to Negro and white oil refinery applicants. Both racial groups were matched on age and education. The spatial (nonverbal) test was found to fail more Negroes than traditional tests. This supports the idea that the nonverbal approach may not be the most nondiscriminatory.

Farr, O'Leary, Pfeiffer, Goldstein, & Bartlett (1971) attempted to develop learning measures of performance. These writers indicated that "...differential reinforcement of basic ability patterns could result in various minority groups being at different points in the learning curve. Thus, if the Negro's cultural background reinforced a pattern of abilities which differed from that of the white subgroup, he would not be at the same point on the learning curve as his 'equally capable' white counterpart" (p. 116). This differential reinforcement puts some members of minority groups at a lower point on the learning curve than members of the majority group. Current tests, then, are poor predictors of ability to learn. A learning situation as a measure of ability should not depend on past learning. Farr et al. used miniature learning tasks, derived directly from the criterion, to predict criterion performance. Forty-six white and 48 Negro college students were used as subjects. The learning tests consisted of: a paired associates task, a concept learning task, and a principle learning task. In addition, the Wonderlic, a vocabulary test, an addition test, and a digit span test were also administered. The criterion was measured performance in a programmed instructural text in statistics. The results demonstrated that whites exhibited more gain over trials than blacks in the concept learning and principles learning tasks. There was no difference in the performance of Negroes and whites over trials on the paired associates learning. With regard to the criterion, the best overall predictors were the Wonderlic and the Vocabulary tests. The principles learning task and the concept learning task also predicted some of the criterion variables. The paired associates task was unrelated to the criterion.

Performance Comparative Studies

A number of studies have also comparatively examined the performance of different groups.

Fox, Taylor, and Caylor (1969) used training tasks to compare low, middle, and high aptitude subjects on: (a) visual monitoring, (b) rifle assembly, (c) missile preparation, (d) phonetic alphabet learning, (e) map plotting, and (f) combat plotting. The low aptitude group needed 2-4 times more training time, 2-5 times more training trials, and 2-6 times more prompting than the middle and high aptitude subjects. The middle aptitude group performed more like the high aptitude group than the low aptitude group. The authors urged that training programs be designed which account for these individual differences.

Guinn, Tupes, and Alley (1970) examined training performance in groups differing in race, education, and area of the country for several occupational specialties. Differences in training performance were found when the groups were divided on the basis of these variables, but differences were not found for all of the occupational specialties on the three variables of interest.

Grunzke, Guinn, and Stauffer (1970) performed a followup study of 26,315 low aptitude (category IV) men accepted into the Air Force. After a comparison with normal enlistees, it was found that low aptitude men:

1. were less likely to complete basic training
2. had more unsuitable discharges
3. were less likely to attain required skill levels.

Van Matre and Harrigan (1970) compared the performance of 54 marginally qualified electronic technicians with 51 well-qualified electronic technicians who underwent training. Performance ratings were obtained in the Fleet on all 105 subjects after they were on the job for 24 months. The low aptitude group was rated similar to the normal group, with none lower than average. Generally, though, the higher aptitude group was rated as more capable in "trouble shooting" and "use of test equipment."

In another study, Van Matre and Steineman (1966) trained 26 low aptitude men in an electronics technician course in an abbreviated period of time. The men were only taught skills considered to be more immediately useful on the job. This experimental group was compared, in a six month followup, with 24 conventionally trained, non-low aptitude subjects. The results demonstrated that the performance of the low aptitude group was adequate and not significantly different from the conventional group.

Van Matre (1971) developed an instrument reading training course for low aptitude (category IV) personnel. Instrument reading is an ability required in many Fleet jobs in which category IV's take part. One-hundred and eighty-eight low aptitude subjects were trained by a variety of methods: classroom lecture, on-the-job training, modified classroom, and independent self study with workbooks. Evaluations were in the form of written criterion tests taken before and after training. The results of this effort demonstrated that group IV men could be trained to the level of experienced non-group IV men. The most effective training method was self study with workbooks.

Hooprich (1968) sought to determine the appropriateness of commisaryman training for category IV personnel. His conclusions, based on two successive studies, were:

1. 31 of 35 category IV's successfully completed training
2. the grades of category IV's, though, were significantly lower than the grades of noncategory IV's
3. the low aptitude men needed to devote more outside time to study and they required more time from instructors to meet criteria
4. the differences between the category IV's and the higher aptitude men were most evident on paper and pencil tests, and least evident on actual performance tests
5. AFQT scores did not predict school performance.

Standlee and Saylor (1969) performed a similar study with equipment operators and obtained identical results.

Rohwer, Ammon, Suzuki, and Levin (1971) worked with 288 elementary school children divided equally over kindergarten, first, and third grades, and also over middle to high socioeconomic white and low socioeconomic black groups. All subjects were given the Peabody Picture Vocabulary Test, the Raven Progressive Matrices, and a Paired-Associates Test. The results showed differences between whites and blacks at all grade levels on the Peabody Picture Vocabulary Test and the Raven Progressive Matrices, thus supporting the conclusion that differences in school achievement may be due to a learning deficiency. The whites, though, were significantly better than blacks only at the kindergarten level on the Paired-Associates test. These authors suggested that the other tests may require "... the mastery of sets of formal conventions (e.g., numbers and categories) created by cultural consensus that may be more readily available to, or more valued by, one population than by another" (p. 13).

Motivational Considerations

One of the critiques leveled at Jensen's (1969) thesis that test score differences between Negroes and Caucasions cannot be accounted for by the environment alone is that motivational differences between the races was not investigated. Jensen discussed motivation in his monograph, but not in as sophisticated a manner as he discussed genetics and environment. Several theories of motivation exist which may account for some of the variation in test scores between races that Jensen attributes to genetic endowment.

Rotter (1966) conceives the effect of reinforcement on behavior as dependent upon whether the person perceives a causal relation between his own behavior and the reward. If not, the reward is attributed to luck and to the control of others. Internal control exists when the subject thinks reinforcement is contingent upon his own behavior, while external control exists when the subject thinks reinforcement is controlled by others or by chance events. In social learning theory, reinforcement increases the expectation that behavior is followed by a reinforcement. Failure of reinforcement extinguishes this expectancy. Children, during development, will thus begin to distinguish causal from noncausal events, and these expectancies will control choice behavior and performance. A person will perceive a reinforcement sequence as not being chance controlled when the proportion of reinforcement is significantly different from 50:50 in a right-wrong situation. When the reinforcements are patterned and when variability is minimal in a task allowing great variability, the reinforcement is perceived as determined by others.

It seems that internal-external control should be considered as antecedent factors in applicant or employee assessment. The only meaningful results would be obtained with subjects who are internally controlled. It is also easily seen that internal-external control can be an important covariate that accounts for differences in intelligence test scores between deprived and nondeprived groups.

In one study investigating the internal-external control concept (Scott & Phelan, 1969), Rotter's Internal-External Control Scale was administered to three groups of subjects. The subjects in all three groups were matched on age, socioeconomic status, and scholastic aptitudes. The results demonstrated that blacks and Mexican Americans demonstrated greater external control than whites. The authors concluded that the externally controlled subjects did not think that there was a relationship between individual effort and reward; therefore, they didn't work unless given external reinforcement (praise, money, etc.).

Battle and Rotter (1963) used Rotter's Internal-External Control Scale to measure external control in several groups of Negroes and whites differing in socioeconomic status. They found that lower class Negroes were higher on external control than lower class whites, middle class Negroes, and middle class whites. Perhaps the perception of limited material opportunities and of powerful external forces produce an external control attitude.

Seeman (1963) and Seeman and Evans (1962) introduced the concept of alienation, which they suggest to be measured by internal-external control. Alienation is a feeling of powerlessness or inability to control outcomes, and it is inversely related to knowledge about a situation. In one study, using reform school subjects, Seeman (1963) sought to determine if poor learning was produced by powerlessness, or if powerlessness comes from poor learning. He found that alienation affected learning about parole information; he therefore concluded that expectancies govern attention and acquisition of knowledge.

Rosenhan (1966) posited that lower class children are more alienated from the environment than middle class children in a middle class school system. Rosenhan then hypothesized that lower class children would be more responsive to praise in a binary choice game than middle class children, and that lower class performance would be more disrupted by disapproval than middle class performance. Rosenhan found an interaction between class and approval-disapproval. Over trials, the lower class approval group started by performing lower than the middle class groups, but finished higher than both middle class groups. The performance of the lower class disapproval group, though, was disrupted in that they remained at a low level across trials. Also, performance of both middle class groups remained the same throughout the trials. From these results, Rosenhan concluded that the lower class child is unfamiliar with middle class institutions, and therefore more alienated and very responsive to external social reinforcement. The middle class child, though, doesn't need external indices of performance. Continual disapproval, then, can have long term deleterious effects for lower class children.

Atkinson (1966) presented a somewhat more rigorous theory of motivation involving achievement motivation, incentive, and goal expectancy. Atkinson's theory is depicted in the following formula:

$$\text{Motivation} = f(\text{motive} \times \text{expectancy} \times \text{incentive})$$

With nAch (motivation to approach a goal) held constant at 1.00 and with expectancy and incentive equal to .5, then the probability of goal approach is .25 (the highest possible). Atkinson defines incentive as goal attractiveness, and motive as the ability to strive for satisfaction or to accomplish. "The strength of motivation to approach decreases as probability of success increases from .50 to near certainty (P_s = .90), and it also decreases as P_s decreases from .50 to certainty of failure (P_s =.10)" (p. 17).

From the above formulation, it is easily seen that the young, deprived black child will rarely encounter a probability of success of .5 or greater. Because he perceives a certainty of failure, he then lacks the motivation to approach a goal, and therefore he does not perform as well in assessment situations as the nondeprived white child who perceives a higher probability of success.

Several recent studies were performed showing class and race differences in nAch. For instance, Rosen (1959) found nAch to be lower in Negroes from middle and lower social classes than other ethnic and racial groups. Negroes in the upper social classes, though, were high in nAch. Rosen (1956), in another study, found that nAch increased as social class increased in white high school males. Mingione (1965) found among low socioeconomic groups that whites had higher nAch scores ($p < .001$) on the TAT than Negroes. Finally, Shrivasta and Tiwari (1967) observed higher nAch in middle class children than in lower class children.

Katz (1967) more or less integrated certain earlier theories into a coherent two-stage theory of development which has strong implications for assessment. During the first stage (up to two years of age) of development, the child's verbal efforts are normally reinforced by parental approval. Selective approval, on the part of the parents, can develop strong habits of striving for proficiency in the child. During stage two, the parental standards and values of achievement are internalized by the child. "The child's own implicit verbal responses acquire through repeated association with the overt responses of the parents the same power to guide and reinforce the child's own achievement behaviors.... Internationalization doesn't take place until strong externally reinforced achieving habits have developed" (p. 5). Lower class children (including most blacks) are more dependent on others for social reinforcement in academic situations. Lacking internalization, they will avoid achievement situations and concentrate on other situations regarded as more promising. "Lower class Negro children tend to be externally oriented in situations that demand performance. That is, they are likely to be highly dependent on the immediate environment for the setting of standards and the dispensing of rewards" (p. 8). Achievement motives and dependency motives must therefore be accounted for in employee and applicant appraisal programs.

Crandall and his associates (Crandall, Preston, & Rabson, 1960; Crandall, Katkovsky, & Preston, 1962; Katkovsky, Crandall, & Good, 1967) have also attempted an integrated theory of motivation. They found that warm, praising, protective, and supportive parental behaviors fostered child belief in internal control, while dominant, rejecting, and critical parental behaviors were negatively associated with internal control. Also, children whose achievement efforts were rewarded as young children later come to value achievement activities as sources of satisfaction.

Hess and Shipman (1965) presented a very interesting development formulation which goes further than the previously mentioned conceptualizations in explaining the differences between Negro and white test scores. These writers indicated that cognitive growth is "... fostered in family control systems which offer and permit a wide range of alternatives of action and thought and that such growth is constricted by systems of control which offer predetermined solutions and few alternatives for consideration and choice" (p. 870). In the deprived family context, the

parent-child control system..." restricts the number and kind of alternatives for action and thought that are open to the child: such constriction precludes a tendency for the child to reflect, to consider, and choose among alternatives for speech and action. It develops modes for dealing with stimuli and with problems which are impulsive rather than reflective, which deal with the immediate rather than the future, and which are disconnected rather than sequential" (p. 870-871). Hess and Shipman concluded that the family shapes the modes of communication in the child which, in turn, shape his thought and problem solving style.

Another motivational conception was presented by Cole and Bruner (1971). These writers dismissed the idea that one group is culturally superior to another. They indicated that what really exists is cultural differences. Persons who are classed as culturally deprived are not presented with situations in which they can demonstrate their skills. Instead, they are continually measured via the middle class culture. This culture is not coincident with their experience and is, accordingly, irrelevant to testing situations. Cole and Bruner conclude that "...cultural differences reside more in differences in the situations to which different cultural groups apply their skills than to differences in the skills possessed by the groups in question....cultural deprivation represents a special case of cultural difference that arises when an individual is faced with demands to perform in a manner inconsistent with his past (cultural) experiences" (p. 874).

In the final two studies reported in this section, the first (Friedrichs, Hertz, Moynahan, Simpson, Arnold, Christy, Cooper, & Stevenson, 1971) used five year old middle and upper middle class children as subjects, while the second study (Stevenson, Williams, & Coleman, 1971) used lower class disadvantaged (mostly Negro) children as subjects. All subjects in both studies were given the same eight learning tasks. Some of the learning tasks were cognitive, while others were associative, therefore providing a good test of Jensen's (1969) thesis. The resulting overall pattern of correlations for the learning tasks was very similar across groups. The authors concluded that "...there is little utility in positing differences between the two groups in the operation of associative and cognitive learning abilities" (p.183).

In summation, these positions reveal with devastating clarity why Jensen (1969) seems to be incorrect in his ascription of heredity as partially causing Negro-white test score differences. This does not mean, though, that the cognitive styles of deprived and nondeprived persons are the same, as Hess and Shipman (1965), and Katz (1967) have so aptly pointed out. Clearly, in training program development, applicant appraisal, and employee development, these differences in cognitive style and motivation must be accounted for and taken into consideration so that the potential of the human resources in our society can be maximized. Motivation and cognitive style variables should therefore be controlled in any study in which racial test score differences are considered. If these factors are not assessed, regardless of how many other variables are included, one's conclusions are apt to be misleading or erroneous.

Discussion

The literature reviewed suggests that for most testing situations, differential validity will exist across subcultures. Accordingly, differential testing methods and different regression equations are indicated for each subculture.

One solution may rest in the "mini aptitude" test approach described in subsequent sections of this report. Here, advantage is take of: (1) the concrete type thinking attributed by others to culturally deprived individuals, (2) the motivational aspects since the paper-and-pencil approach is avoided, and (3) the minimization of culturally loaded content.

One problem is that not all minority group members have led a deprived existence. How can we determine if a Negro or a Spanish American is deprived? Certainly skin color and other physiognomic characteristics are associated with deprivation, but they do not tell us with certainty that any one individual is disadvantaged. A solution is to use a well-constructed Biographical Information Blank (BIB) containing questions related to cultural exposure. The Negro child who owned 200 books, and who visited museums as a child is definitely less deprived than the Negro child who rarely saw a newspaper, much less a book. Many of the differential prediction studies may have "missed the mark," because they have used race as a moderator variable rather than cultural deprivation. The only reason some of these differential prediction studies have been successful is because race is correlated with cultural deprivation. Naturally, more Negroes will be in the culturally deprived group, but some whites will be there, too.

Kimble (1971) is one of the few investigators who have successfully developed a cultural deprivation scale. Kimble's scale was based on three classes of variables: (1) amount, (2) variety, and (3) organization of stimulation. This scale was administered along with the School and College Abilities Test (SCAT) to 200 students in remedial classes at a junior college. Kimble found that the subscales of his cultural deprivation scale correlated between -.32 and -.44 with the SCAT.

Purpose of Present Work

The primary purpose of the present work is to investigate, in the Navy context, assessment methods, techniques, and procedures which are free from the biases ascribed to more traditional testing approaches. The study does not focus on a test or tests which possess equal predictive validity for both high and low aptitude personnel. Rather, the assumption is made that the normal Navy testing vehicles (GCT + ARI + MECH + CLER) are adequate for persons possessing high aptitude as measured by these methods. The test results of persons who achieve high

scores on these instruments have not been affected to the extent that their progress in the military will be debilitated in any way. On the other hand, the results of persons who score poorly on these tests may be unduly affected by the factors discussed in earlier sections of this report. Accordingly, the present study attempts to develop predictive procedures which will identify low aptitude personnel (as measured by the usual Navy tests) who can perform adequately on the job.

The underlying working hypothesis is that persons who exhibit the ability to learn sample aspects of a Navy job will be able to learn the total job, provided that they are given proper on-the-job training. A similar concept has been previously developed by Jensen (1969) and by Farr, O'Leary, Pfeiffer, Goldstein, and Bartlett (1971).

The specific research steps include:

1. development of a sample of miniaturized job learning situations (tests) for low aptitude personnel in the machinist mate (MM) rating. These miniaturized job learning situations are called training and evaluation situations in subsequent sections of this report

2. administration of these tests to a sample of low aptitude black and white persons and assigning these persons to probable successful and probable unsuccessful groups on the basis of their test scores

3. assignment of all persons sampled to Fleet jobs in the machinist mate rate

4. followup to determine the degree of on-the-job success experienced by all persons in the sample.

To date, steps 1, 2, and 3 have been completed. The methods, procedures, and results of these steps are reported in subsequent sections of this report.

Followup in the Fleet (step 4) has not yet been accomplished. Followup studies are planned after the tested individuals have been on their assigned jobs 6 months, 12 months, and 24 months.

CHAPTER II

METHODS

After discussion with persons occupying a number of desks concerned with Navy personnel and training methods and procedures, the machinist mate job was adopted as at least a logical starting place for a study such as that which is here involved. The machinist mate rate in the Navy involves performance of tasks which are largely nonverbal in character. The thinking processes involved in these tasks are largely concrete (as opposed to abstract) in nature. Moreover, this rate is one in which there is not a large number of blacks. Nonetheless, it is a rate which should be attractive to most recruits since it is adequately high on the informal prestige scale for various Navy jobs and because it offers the potential for learning skills which can lead to post Navy employment.

The normal entry into the machinist mate career field in the Navy is through the Navy "A" school for machinist mates. This school involves training in the fundamental skills and knowledges required for performance at an entry level in this rate.

Subjects

The subjects were Navy recuits who were identified after initial Navy testing by the Recruit Training Command at the Great Lakes Naval Training Center. As a basic requirement for participation, a subject had to have "failed"* the entry tests for the Machinist Mate School. Ninety-nine recruits were so identified. Fifty of these recruits were white and 49 were black. Virtually all of the recruits were between 19 and 20 years of age.

Miniature Training and Evaluation Situations

As a first step in the construction of the miniature training and evaluation situations, the machinist mate section of NAVPERS 18068A was consulted. Those practical behaviors required for advancement to level E-4 were extracted. Several of these behaviors were combined because they were of a similar nature. The next step involved a meeting with five Master Chief Machinist Mates and one warrant officer at the Great Lakes Naval Training Center. During this meeting, a final list of behaviors, which were adequately representative of the most frequently performed or critical tasks of the journeyman level machinist mate, were agreed on.

---

* A recruit must exceed a combined GCT, ARI, and MECH score of 158 to be eligible for entry to the Machinist Mate Class A School.

The behaviors so identified were:

1. ability to identify and use hand tools common to job
2. ability to perform maintenance and to read meters and gauges accurately when under some degree of distraction, vigilance, or when attention sharing is involved
3. ability to make simple repairs in pressure lines
4. ability to perform simple troubleshooting and systems analysis in pressure systems
5. ability to operate equipment common to rate
6. ability to assemble and disassemble common high failure frequency items

These tasks formed the basis for the miniature training and evaluation situations. Several Master Chief Machinist Mates then served as technical consultants during the actual lesson construction phase of the project. This procedure resulted in the construction of six miniature training and evaluation lessons. These training and evaluation situations reflected samples of the most critical and/or frequently performed behaviors of the journeyman machinist mate.

Each training and evaluation situation contained two segments. The first segment, a training phase, usually involved a "show and tell" learning situation of 15 to 30 minutes. All of these lessons were completely devoid of reading and writing requirements.* The sections which follow describe each of the training and evaluation situations in detail.

Equipment Use and Nomenclature

In the equipment use and nomenclature mini job sample situation, the objective was to determine if the recruits could learn the names and uses of all of the equipment and material involved in making-breaking a flange. It was assumed that if a sailor could learn the names and uses of the tools and materials involved in this situation, he would also be able to learn the names and uses of other equipment used on the job. The materials used in performing this repair task are: (a) bolts, (b) nuts, (c) flanges,

* Some of the tests involved the ability to tell time and the ability to read numbers. These are considered to be preliterate requirements.

(d) drift pins, (e) gasket knife, (f) gasket material, (g) graphite grease, (h) ball pean hammer, and (i) box end wrenches. After a tape recorded introduction, the instructor demonstrated how to make and break a flange. During this demonstration, the use and name of each piece of equipment was discussed by the instructor. For example, before using a particular tool, the instructor would hold up the tool and say "This is a _______." He would then pass the tool among the students for their inspection. When the tool was returned to the instructor, he would demonstrate its use in the flange repair situation. Upon completion of the demonstration, a 25 question true-false test was administered to the recruits. For each item in this test, the instructor held up an object and ascribed a name or use to it. The recruits then indicated whether the name or use given by the instructor was true or false by encircling either the word "true" or the word "false" next to the item number on their answer sheets. Each item was read twice with a 10 second interval between items.

Gasket Cutting and Meter Reading

The gasket cutting and meter reading training and evaluation situation was designed to investigate ability to learn a maintenance task and to perform when some degree of distraction or attention sharing is involved. This situation was also designed to sample the vigilance situation in which the machinist mate, on the job, must monitor the states of various equipment systems while he performs other tasks. After a tape recorded introduction, the subjects were taught, through demonstration, how to make a gasket using a flange, a ball pean hammer, asbestos gasket material, and some bolts. In this demonstration, the gasket material was placed over the flange so that the face of the flange was entirely covered. Then, using the round end of the ball pean hammer, one bolt hole was tapped out. Only light taps were suggested, because heavy hammering on the flange would eventually damage it. Next, a bolt was place in the bolt hole. The next step was to tap out a bolt hole, exactly opposite the first one. This was done so that the gasket would not shift over the flange. The remaining bolt holes were then tapped out in any order. The inner circumference was then tapped out, again using the round end of the ball pean hammer. The flat end of the ball pean hammer was then used to tap out the gasket material remaining over the outside edge of the flange. Finally, the tapped out pieces of gasket material were removed from the flange and the gasket.

Upon completion of the lesson, the subjects were given a 10 minute gasket making practice session. During the practice sessions, the instructor circulated among the recuits and assisted them as required.

Next, the subjects were taught how to read a pressure meter and: (1) how to log the time at which the pressure deviated from the normal, and (2) whether the pressure should be adjusted to go up or down (relative to a given nominal value). A large clock with a sweep hand was placed in front of the testing room. The recruits read the time from this clock.

The tests for both gasket cutting and meter reading were administered together. That is, for a ten-minute period, the subject had to observe and record from a meter while he constructed a gasket. The meter which each student read was placed at his individual work station. Each meter was individually driven so that there was no possibility for a subject who noticed a system out-of-tolerance condition at his station to cue a recruit at another station of an out-of-tolerance condition at the second station. However, the signal presentations to all stations were equated for number, direction, and magnitude of deviation. This combination performance testing introduced the required attention sharing component into the criterion situation.

The gasket making was scored through a checklist which was completed by the instructor as the recruit performed the task. The scoring checklist included items on adherence to correct procedures, care and use of tools, adherence to safety precautions, and adequacy of the final gasket. The meter reading aspect was scored on the basis of: (1) number of out-of-tolerance conditions correctly noted, (2) correctness of indication of needed pressure adjustment (up or down) to restore system to nominal, and (3) precision of log entry for time of deviation.

The scoring checklist for the gasket cutting is included in Appendix A to this report. Figure 2-1 shows the black instructor demonstrating the gasket making procedure. Figure 2-2 shows one of the recruits and the instructor during the gasket making practice.

Figure 2-1. Demonstration of gasket construction procedure.

Figure 2-2. Examinee practicing gasket cutting.

Trouble Shooting

The goal of the trouble shooting mini job situation was to test the recruit's ability to learn to perform elementary system analysis and trouble shooting on a hydraulic pressure system. A simulated pressure system was used as the apparatus for both the training and testing aspects. After a taped introduction, the subjects were taught how the pressure system operates. Essentially, the apparatus consisted of a set of color coded and interchangeable gears which were so interconnected that a simulated pumping system was driven. A schematic representation of the system is shown in Figure 2-3. Within the simulated pumping system, a set of valves controlled the flow. Accordingly, to diagnose a fault in the system, the recruit needed to understand such elementary relationships as: (1) the effects of gear size on pump speed/rate of flow, (2) the effects of direction of gear rotation on flow, (3) how large and small gears can intermesh to produce changes in output rate, and (4) the effects of valve and pump function on system operation.

A series of simulated light indicators was used to indicate the adequacy of the flow at various portions of the simulated system. The task of the subject was to observe the light indicator, determine whether any out of tolerance conditions existed, state the cause of the condition, and what should be done to remedy the out of tolerance condition, if any.

After the operation of the simulated system had been explained, various malfunction situations were presented, and the recruits were taught what locations in the system needed adjustment in order to correct the problem and the cause of the malfunction. After the training, the subjects were presented with a number of practice trouble shooting problems. A typical problem was:

1. low pressure light "on"
2. pump speed indicator on "low"
3. bypass valve "open"
4. Say: "THERE IS NO LEAK IN THE SYSTEM. WHAT HAS TO BE DONE TO FIX THIS?"

For the practice problems, both the correct answer and the reason for it being correct were discussed.

In the test situation, 12 problems were involved. The subjects circled one or two of nine numbers corresponding to nine possible malfunction causes. Figure 2-4 shows one of the instructors explaining a detail in the pressure system.

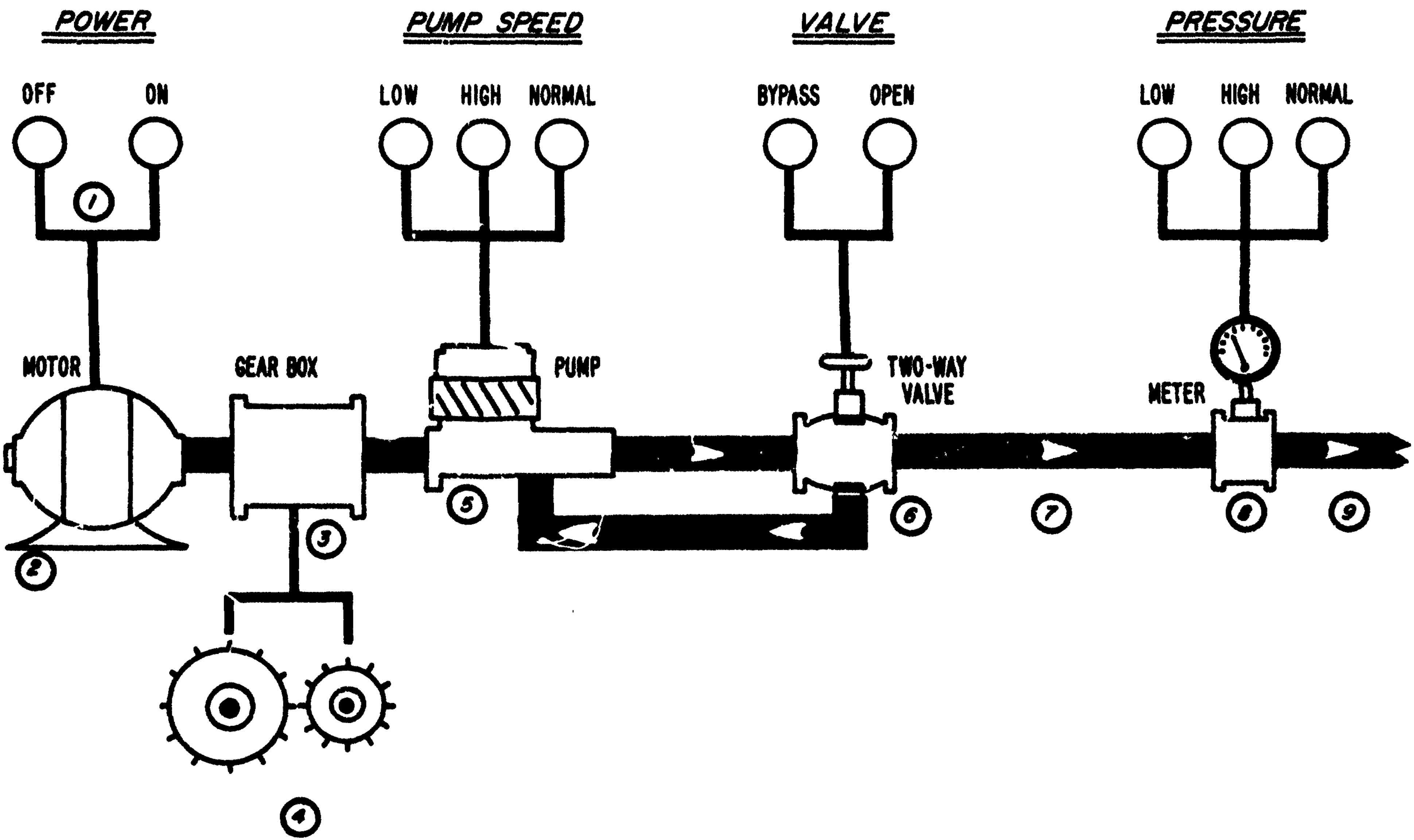

Figure 2-3. Schematic representation of pump system for trouble shooting test.

Figure 2-4. Instructor explaining function of simulated pressure system.

Equipment Operation

In the equipment operation mini job learning situation, the subjects were taught how to start up and shut down a motor and pump apparatus. The students were required to learn a 33 step procedure, including several safety precautions. Each class member was then given an opportunity to practice starting up and shutting down the apparatus. The equipment operation procedure (somewhat abbreviated) includes the following:

1. removal of oil rags from under motor
2. checking oil and gasoline levels with dipsticks
3. adding fuel and oil from fuel and oil storage tanks
4. replacing lids on oil and fuel tanks
5. checking to determine if valve is in bypass position
6. plugging in battery socket
7. turning ignition switch
8. turning bypass valve on, then off
9. turning off ignition
10. cleanup of oil and fuel spills

After practice, a checklist type performance test was administered. Scoring was completed while the recruit performed and was based on adherence to correct procedures and observance of safety precautions. All relevant materials for the equipment operation test are listed in Appendix of this report.

Figure 2-5 shows a group being instructed in equipment operation. Figure 2-6 shows a sailor practicing on the equipment, and Figure 2-7 shows a subject being tested.

Figure 2-5. Instruction on the equipment operation apparatus.

Figure 2-6. Subject practicing on equipment operation apparatus.

Figure 2-7. Subject being tested on equipment operation apparatus.

Assembly

In the assembly miniature learning situation, the recruits were taught how to assemble a gate valve from its component parts. First, a demonstration of the correct assembly procedure was presented. The correct assembly procedure includes:

1. screwing packing nut to top of stem
2. screwing gate to bottom of stem
3. winding gate all the way up the stem
4. screwing gate and stem assembly onto body of valve
5. screwing handle onto top of stem with handle nut
6. screwing on two 3/4" nipples
7. checking to determine if parts are fitted tightly together

This demonstration was followed by a short practice session in which the students were allowed to assemble the valve themselves. The instructor circulated among the students during this practice session and assisted each one, as required. After the practice, each subject was individually tested on his ability to assemble the valve. Again, scoring was through the checklist procedure.

Figure 2-8 shows the instructor presenting the correct valve assembly procedure. Figure 2-9 presents the practice session, and Figure 2-10 shows a student beginning the valve assembly test. All relevant test materials for the valve assembly task can be found in Appendix A of this report.

Figure 2-8. Instructor teaching valve assembly procedure.

Figure 2-9. Two students practicing valve assembly.

Figure 2-10. Student beginning valve assembly test.

Pass-Fail

A subject was judged to "pass" the miniature evaluation battery if he scored "average" or better on the trouble shooting test and "average" or better on two of the five remaining tests. Such subjects were assigned to a "probable successful" group and assigned to a ship in the Fleet for work in the machinist mate rate. Subjects not achieving this level were assigned to a "probable fail" group and were similarly assigned. Only Applied Psychological Services is aware of which recruits have been placed in each group.

The logic for the choice of cut scores was that the trouble shooting mini job learning and test situations were largely cognitive in nature, while the remaining situations largely involved learning manipulative procedures. Thus, those in the probably successful group exhibited some cognitive as well as manipulative/procedural skill learning ability in the mini job learning situation.

The scores of each recruit in the sample on the GCT, ARI, MECH, CLER, and SP tests of the normal Navy classification battery were provided by the Recruit Classification Center at the Great Lakes Naval Training Center. These data are presented in Chapter III of the present report and indicate fairly close equivalence between the white and the black groups on these tests.

Instructors and Test Administrators

Two instructors/test administrators managed each training and evaluation session. One instructor/test administrator was a 36 year old retired, black Navy Chief Petty Officer. At the time of the present program, he was employed by the North Chicago school system, which generously granted him leave time for the purposes of this program.

The second instructor/test administrator was a 29 year old white psychologist. The assigned duties of the black Chief were:

1. taping the introduction to each lesson
2. conducting the training segment of each lesson
3. administering the tests to black recruits

This instructor was thoroughly trained in the content he was to present, the teaching methods he was to employ, and the test procedures prior to implementation of the present program.

The duties of the white psychologist were:

1. organization and direction of the mini job sample learning and testing program
2. training the black instructor/test administrator
3. assisting the black instructor/test administrator in his lesson presentations
4. testing the white recruits

A black instructor/test administrator was used, because it seemed that the black recruits would be more motivated when the proctor was a member of their own race. Recent research (Cole & Bruner, 1971) has shown that low achievement black students will perform at a more optimal level for a black proctor than for a white proctor. In addition, the education, personality, verbal inflection, and method of treating the subjects of the black instructor/test administrator were of such a nature that he could easily be identified with and understood by the black recruits.

Setting

All of the mini job sample learning and testing sessions were held in a large classroom provided by the Machinist Mate School, Naval Training Center, Great Lakes. This classroom was equipped with 12 student desks and six worktables of various sizes. The lighting, temperature, ventilation, space, and privacy were considered optimal for this study.

All research sessions began at 0720 hours. A one-hour lunch break was allowed between 1130 and 1230 hours. Most sessions were completed by 1500 hours. Table 2-1 shows the approximate amount of training, practice, and testing time for each mini job sample training and testing situation. The situations required from 25 to 65 minutes of total training, practice, and testing time. These times are well within the acceptable range of testing times currently used in educational and military settings.

One half of the subjects were processed between 3 December and 16 December 1971. The remaining half were trained and evaluated between 10 January 1972 and 21 January 1972.

Table 2-1

Approximate Training, Practice, and Testing Time for Each Mini Testing Situation (minutes)

	Training Time	Group Practice Time	Individual Practice Time	Group Testing Time	Individual Testing Time	Total
Equipment Use and Nomenclature	15	0	0	10	0	25
Gasket Cutting/ Meter Reading	15	15-25	0	10-20	0	60
Troubleshooting	35	15	0	15	0	65
Assembly	10	10	0	0	3*	32
Equipment Operation	15	0	5*	0	3*	47
Total	90	50	20	45	24	229

* Based on a group of eight students with two proctors present

Questionnaire

A 36 item personal background questionnaire was constructed by Applied Psychological Services to measure various facets of cultural deprivation including: (a) need achievement, (b) home environment, (c) school environment, and (d) other demographic variables. This questionnaire was administered to all the recruits in the mini job learning situation sample. It was considered that these cultural factors could conceivably moderate learning ability to the extent that the correlations between the miniature evaluation test scores and the ultimate performance criteria would be lowered.

In addition, Applied Psychological Services administered the same questionnaire to a control group of Machinist Mate A School recruits, who had met or surpassed the screening criteria for that school.

Interview

After completion of the final mini job sample learning and testing situation, each recruit was interviewed by one of the instructor/test administrators in order to obtain reactions to the entire training and testing program. Basically, the subjects were asked to compare the tests and training they received in the present program with other types of tests and training they received in the past. All subjects were encouraged to respond freely and openly to the interviewer. The interview questions were constructed in a manner allowing quantitative and qualitative analysis of the interviewee responses.

CHAPTER III

RESULTS

Intercorrelations

The correlations among the scores of recruits on the six mini job sample learning tests were determined, along with the correlations among mini job sample learning scores and scores on the Navy classification tests. The intercorrelation matrices are shown for each racial group in Tables 3-1 and 3-2. Examination of these matrices indicates, rather conclusively, that the six miniature job learning tests are measuring factors which are quite unique from the factors measured by the Navy classification tests. These results are probably due to the fact that the Navy classification tests are heavily loaded in verbal and cognitive components, while the miniature job learning tests emphasize perception and performance. In addition, the correlations demonstrate that the miniature job learning tests are relatively independent from each other. Finally, we note that the single correlation coefficient of any substantial magnitude, the correlation between the mechanical test of the Navy classification battery and the gasket cutting mini job learning test (for the white group only), might have been anticipated on the basis of the mechanical aspects of both of these tests.

Means and Standard Deviations

The means and standard deviations of the recruits sampled on the Navy classification tests and on the miniature learning tests were also computed by racial group. The results are presented in Table 3-3. For the Navy classification tests, although the means and standard deviations of the white group are higher for all the tests, none of the mean differences are statistically significant ("t" test). However, there is a small but statistically significant variance difference ($p < .05$, Fmax test) for the mechanical and the clerical tests of the Navy classification tests. For both of these tests, the variance of the white group was greater than that of the black group.

This result supports a contention that the attempt to match the two groups on the basis of their Navy classification test scores was moderately successful.

Of considerably more importance is the finding that for the miniature job sample tests there are no statistically significant differences between racial groups. This holds for both the mean differences and the variance differences. Accordingly, it seems that the mini job learning situational tests can be held to possess minimum cultural bias.

Preceding page blank

Table 3-1

Pearson Product Moment Correlations Between the GCT, ARI, MECH, CLER, SP, Equipment Use and Nomenclature (EUN), Gasket Cutting (GC), Meter Reading (MR), Trouble Shooting (T), Equipment Operation (EO), and Assembly (A) Tests for 50, Low Aptitude, White, Naval Recruits

	Test									
	ARI	MECH	CLER	SP	EUN	GC	MR	T	EO	A
GCT	.52	-.15	.38	.07	.15	-.16	-.05	-.10	-.07	.13
ARI		.13	.41	.03	.11	.01	.25	.13	.14	.08
MECH			.20	.30	.21	.64	.13	.25	.19	.01
CLER				-.10	.15	.05	.00	-.04	-.15	-.06
SP					.22	.26	-.08	.07	.24	.03
EUN						.18	.14	-.06	-.06	-.09
GC							.09	.29	.23	.01
MR								.27	-.09	-.17
T									.38	.16
EO										.08

Table 3-2

Pearson Product Moment Correlations Between the GCT, ARI, MECH, CLER, SP, Equipment Use and Nomenclature (EUN), Gasket Cutting (GC), Meter Reading (MR), Trouble Shooting (T), Equipment Operation (EO), and Assembly (A) Tests for 49, Low Aptitude, Black, Naval Recruits

	Test									
	ARI	MECH	CLER	SP	EUN	GC	MR	T	EO	A
GCT	.20	.16	.29	.22	.09	-.25	.20	.14	.09	.10
ARI		.37	.24	.36	-.19	.23	.06	.25	.02	.16
MECH			.01	.31	-.11	-.04	.23	.16	-.04	-.02
CLER				.36	.14	-.21	.15	.11	.06	.05
SP					.05	-.14	.13	.27	-.09	.19
EUN						-.30	-.09	-.23	.13	.04
GC							-.20	-.08	.02	.19
MR								.31	.14	-.17
T									.21	.19
EO										.09

Table 3-3

Means and Standard Deviations on the GCT, ARI, MECH, CLER, SP, Equipment Use and Nomenclature (EUN), Gasket Cutting (GC), Meter Reading (MR), Trouble Shooting (T), Equipment Operation (FO), and Assembly (A) Tests for 50 White and 49 Black, Low Aptitude, Naval Recruits

	White		Black	
Test	Mean	S.D.	Mean	S.D.
GCT	39.58	7.33	38.29	6.52
ARI	42.56	5.83	41.24	4.27
MECH	42.16	6.22	40.73	4.54
CLER	51.90	12.27	51.82	8.68
SP	43.44	5.30	41.51	6.00
EUN	21.16	3.17	20.59	2.60
GC	14.50	2.78	14.31	2.78
MR	10.92	4.26	10.78	4.34
T	14.60	5.64	13.82	5.56
EO	59.78	7.02	57.88	6.56
A	22.84	3.61	23.29	3.25

"Probably Acceptable" - "Probably Unacceptable"

As stated previously, a recruit who scored at or above the group mean on the miniature job learning trouble shooting test and who scored similarly on any two of the remaining five miniature job learning tests was placed in a group for which "probably acceptable" Fleet performance is predicted. It is anticipated that the members of this group will demonstrate at least minimally satisfactory progress on the machinist mate job in the Fleet. On this basis, 60 per cent of the white recruits in our sample (n = 30) and 53 per cent of the black subjects (n = 26) were placed in the "probably acceptable" category. A chi-square analysis was performed on these data. The results of the chi-square analysis indicated that race was not significantly associated with "passing" or "failing" the miniature evaluation battery. This result was expected, inasmuch as the mean difference across race for each test separately was not statistically significant.

Reliability

The test administration procedures for the equipment operation, assembly, and gasket cutting tests permitted an analysis of interrater reliability. For these tests, a sample of the recruits was scored independently by both the black instructor/test administrator and by his white counterpart. The separate scores, so determined, were compared. The results of this interrater reliability analysis are shown in Table 3-4. Examination of Table 3-4 shows that the interrater reliability coefficients were acceptably high for performance tests of this type. In addition, the means and standard deviations across procters were almost identical. This suggests that scoring methods for the procedural tests were sufficiently objective to allow a reliable total score estimate.

Table 3-4

Means, Standard Deviations, and Interrater Reliability Coefficients for a Black Test Administrator and a White Test Administrator Scoring the Gasket Fabrication (GC), Equipment Operation (EO), and Assembly Tests (A)

	Test					
	GC		EO		A	
Administrator	White	Black	White	Black	White	Black
n	32		39		39	
Mean	14.41	14.72	56.03	55.69	22.05	22.26
S.D.	2.23	3.10	7.49	7.46	3.21	3.33
r	.75		.97		.96	

The Interview

After completion of his learning and evaluation session, each examinee was interviewed by the instructor/test administrator of the race of the examinee. The four questions involved in each interview were:

1. How would you compare the training portion of this program with other training programs you have encountered in the past? Was it better, worse, or about the same? Why?

2. How would you compare the tests you took today with the more traditional paper and pencil tests you have taken in school? Are they better, worse, or about the same? Why?

3 Did you enjoy participating in today's program? Why?

4. Please tell me any other impressions or thoughts you have about the training program.

Seventy-eight per cent of the recruits thought the training portion of the program was better than the training they had received in other programs. Twenty per cent of the subjects thought the training portion of the program was the same as the training they had received in other training programs. Only one per cent of the subjects, though, thought the training they received was worse than that of other training programs. These results support the emphasis placed on performance during the learning sessions rather than on reading and writing. Some representative subject responses to this question were:

1. The training was easy because it doesn't involve reading. Never took one (training program) like it before. I'm at the fourth or fifth grade in reading.

2. Had someone explain it to you step by step. Gave you time to do it.

3. Got more out of it. You learn more.

4. Mechanical stuff is my type of work. It's my thing. I dig working with my hands.

5. You get more out of putting things together than in looking at pictures and directions. You can see what each part looks like, but you can't in a picture. If you have it in your hand you know what it feels like.

With regard to the second question, 86 per cent of the subjects thought the miniature job learning tests were "better than" paper and pencil tests. Twelve per cent of the subjects thought the program tests were the "same as" paper and pencil tests. Again, only one per cent thought the program tests were "worse than" paper and pencil tests. These results support the use of performance oriented tests which require little or no reading. Sample interview responses to this item are:

1. You don't have to do as much writing. In school they made me more tense with more pressure on my mind.

2. Because you see what you're doing. In school tests you have to read it from a book. You have to keep it in your mind. You don't get to see what you're doing.

3. The questions don't drag on. Can't understand the ones in school as well.

4. Just asked the question right after showed how to do it.

5. Because you don't spend the time reading the questions.

6. When you read it yourself, you might not understand the words, but when he read it out you don't get a chance to goof up on the words.

The responses to the final question showed that 98 per cent of the subjects enjoyed the training and evaluation program, while only two per cent did not enjoy it. These results and the results of the first two questions allow the conclusion that "low aptitude" Naval Recruits exhibit an overwhelming attitudinal preference for the learning and evaluation program, as here employed, over the more traditional testing programs.

The Questionnaire

The primary purpose of the cultural deprivation questionnaire in the current investigation is for eventual application as a statistical control. Before the questionnaire is used as a control measure, though, its structure must be established.

Factor Analysis

The questionnaire scores for the entire sample were subjected to a principal components factor analysis with a varimax rotation. Nine factors, accounting for 46 per cent of the predictable variance, were extracted. These were called: self-esteem, environmental stimulation, reading habits, educational attainment, educational initiative, parental interest, monetary deprivation, educational encouragement, and urbanity. The items with heaviest loadings on each of the nine extracted factors are presented in Tables 3-5 through 3-13.

Table 3-5

Items with Highest Loadings on Self-Esteem* Factor

Item	Loading
At this time, what do you think your chances are of successfully advancing in the Navy are?	.757
At this time, what do you think your chances are of successfully passing the tests for one or more of the Navy rating(s) you hope to enter?	.730
Approximately what yearly salary do you think you will be earning ten years from now?	.587
During your past schooling, how would you have done in school if you had done the very best you could?	-.478
Comparing yourself to others you know, how do your decisions seem to stack up in quality?	-.476
How did you compare with other fellows in rate of progress through school?	-.335

*Eigen value = 2.685

Table 3-6

Items with Highest Loadings on Environmental Stimulation* Factor

Item	Loading
How many rooms did your home have when you were a child?	.657
How much education did your mother have?	.582
How much education did your father have?	.565
How many books do you now own?	.420
To how many magazines and periodicals did your family subscribe while you were growing up?	.388
When you were growing up how many books were around the house?	-.635
How do you feel about the achievements of your parents?	-.356

*Eigen value = 2.302

Table 3-7

Items with Highest Loadings on Reading Habits* Factor

Item	Loading
Other than schoolwork, how much reading did you do during your youth?	.767
During your school years, to what extent did you read newspapers?	.608
About how often do you spend an evening at home sitting around and reading?	-.650

*Eigen value = 2.038

Table 3-8

Items with Highest Loadings on Educational Attainment* Factor

Item	Loading
How much education have you had?	.542
How many serious, nonfiction books have you read in the past year, not counting text books?	.480
How did you compare with other fellows in rate of progress through school?	-.567
How difficult was high school work for you?	-.489
How did your parents feel about the marks you made in school?	-.424
During your past schooling, how would you have done in school if you had done the very best you could?	-.318
How often did you seriously consider quitting school?	-.313

* Eigen value = 1.791

Table 3-9

Items with Highest Loadings on Educational Initiative* Factor

Item	Loading
To how many magazines and periodicals did your family subscribe while you were growing up?	.382
How often did you seriously consider quitting school?	-.747
As you grew up how did you feel about school?	-.684
How did you compare with your friends in rate of progress through school?	-.307

* Eigen value = 1.772

Table 3-10

Items with Highest Loadings on Parental Interest* Factor

Item	Loading
While you were in school, how much interest did your parents or guardians appear to take in your school work?	.654
During most of your school years, would you say that your needs were: [(a) well provided for; (b) satisfactorily provided for; (c) somewhat meagerly provided for but tolerable; (d) unsatisfied most of the time; (e) never satisfied]?	.591
How did you feel about the achievements of your parents?	.526
How did your parents feel about the marks you made in school?	.376

* Eigen value = 1.625

Table 3-11

Items with Highest Loadings on Monetary Deprivation* Factor

Item	Loading
When you were a child, did your parents talk or act as though money were a problem?	.652
During your high school years, what was your total family income per month?	.570
During most of your school years, would you say that your needs were: [(a) well provided for; (b) satisfactorily provided for; (c) somewhat meagerly provided for but tolerable; (d) unsatisfied most of the time; (e) never satisfied]?	-.438

* Eigen value = 1.491

Table 3-12

Items with Highest Loadings on Educational Encouragement* Factor

Item	Loading
When you were a child, did your parents compare your school performance (favorably or unfavorably) with that of other children?	.597
When you were a child, did your parents sometimes tell you to stay inside and read more?	.568
As a child, how often did your parents encourage you to read?	.429

*Eigen value = 1.387

Table 3-13

Items with Highest Loadings on Urbanity* Factor

Item	Loading
How would you describe the neighborhood in which you were brought up (degree of crowding)?	.502
The place in which you spent the most time during your early life was a (city size)?	-.671

*Eigen value = 1.367

Significance Tests

In order to test whether or not the "low aptitude" group differed significantly on the cultural deprivation factors from a group which meets the selection standards for the machinist mate "A" school, the cultural deprivation questionnaire was administered to 118 recruits in the "A" school. The factor score means for the "high aptitude" (A school) group and for the "low aptitude" group were calculated. These are presented in Table 3-14. Tests ("t" tests) were conducted between the group mean scores for each factor. The results of these tests are also presented in Table 3-14. For seven of the nine factors, the "low aptitude" group demonstrated significantly more cultural deprivation, as measured by this questionnaire, than the "high aptitude" group. Accordingly, cultural deprivation (as here measured) and "low aptitude" appear to go hand in hand.

Examination of the data in Table 3-14 indicates no statistically significant differences across aptitude groups in the parental interest factor.

In another factor, educational encouragement, the "high aptitude" group seems to be more deprived than the "low aptitude" group. It is quite possible that the "low aptitude" group perceives educational encouragement differently than the "high aptitude" recruits. "Low aptitude" persons may perceive educational encouragement as nagging. The young deprived child then would associate academic pursuits with nagging and parental disfavor, resulting in discouragement of the child's academic efforts. An alternate explanation is that the "low aptitude" group receives more educational encouragement because they tend to perform less well in school. A child who is already performing well does not need as much encouragement as the child who is performing poorly.

Correlation with Tests

One of our hypotheses was that the miniature job evaluation tests would be less contaminated by cultural deprivation factors than the usual Navy qualification tests. The Pearson product moment correlation among the qualification test scores, the miniature job sample tests, and the questionnaire factor scores for the 99 recruits in the "low aptitude" sample are presented in Table 3-15. In general, the correlations between the factor scores and the test scores tend to be low for both the Navy qualification test and the miniature job learning tests. However, only one of the correlations between the mini learning tests and the deprivation factors is statistically significant. On the other hand, seven of the correlations between the Navy qualification tests and the cultural deprivation factors scores are statistically significant. The data in Table 3-15, then, support our contentions that the mini tests are unrelated to cultural factors and are less culturally loaded than the usual Navy qualification tests.

Table 3-14

Means, t-Ratios, and Significance Levels for "High Aptitude" and "Low Aptitude" Recruits on Nine Factors

	High Aptitude (n=118) Mean*	Low Aptitude (n=99) Mean*	t-Ratio
Self-Esteem	-147.40	-5.90	8.51 ($p < .001$)
Environmental Stimulation	-34.58	2.20	3.52 ($p < .002$)
Reading Habits	-21.67	-2.30	2.69 ($p < .01$)
Educational Attainment	-91.64	-0.39	8.71 ($p < .001$)
Educational Initiative	-37.87	1.52	3.00 ($p < .01$)
Parental Interest	6.34	-3.69	0.55 (N.S.)
Monetary Deprivation	-32.37	-13.36	2.17 ($p < .05$)
Educational Encouragement	15.29	-0.18	3.43 ($p < .002$)
Urbanity	-10.86	0.22	3.32 ($p < .002$)

* Lower scores indicate less cultural deprivation.

Table 3-15

Pearson Product Moment Intercorrelations Among the Questionnaire Factor Scores, the Navy Qualification Tests, and the Miniature Job Learning Tests (N = 99)

Factor	Test										
	GCT	ART	MECH	CLER	SP	EUN	GC	MR	T	EO	A
Self-Esteem	.13	.22*	.06	.17	-.14	-.04	.10	-.06	.00	-.10	.03
Environmental Stimulation	.04	-.02	.00	.18	-.05	.05	.09	.04	-.02	.05	.05
Reading Habits	-.14	-.11	.04	-.17	.10	-.01	.02	.00	-.16	.12	-.09
Educational Attainment	.05	.17	.00	.23*	.00	.04	.10	-.06	.08	-.04	.04
Educational Initiative	-.05	-.02	.00	.13	-.08	.02	.01	-.09	.07	-.15	.04
Parental Interest	.00	-.08	-.07	-.17	.15	-.02	-.10	.03	.01	.16	-.10
Monetary Deprivation	.27**	.22*	-.05	.26**	.04	.09	.00	.03	.02	.00	.06
Educational Encouragement	-.01	-.07	.02	-.10	.18	-.02	-.08	.10	.09	-.01	.03
Urbanity	.00	.29**	.25*	-.10	.16	-.02	.21*	.12	.15	-.02	-.05

** $p < .01$
* $p < .05$

CHAPTER IV

CRITERION DEVELOPMENT

Quite obviously, the demonstrated reliability, freedom from cultural bias, equivalence for both white and black groups, and the like represent necessary but not sufficient ingredients for an assessment approach. The predictive validity of the miniature job learning tests remains to be demonstrated. To this end, each of the recruits in our "low aptitude" sample will be followed up after he has served 6 months, 12 months, and 24 months in the machinist mate rate. The longitudinal followups will be based on a criterion referenced test approach, supplemented by supervisor ratings and supervisory interview data. We place little, if any, credence in the supervisory ratings and interview data since it is known that supervisors often rate black job incumbents lower than whites. Flaugher, Campbell, and Pike (1969) found that white supervisors rated a group of Negro incumbents one-half a standard deviation lower than Negro supervisors. Negro supervisors, though, did not rate white incumbents higher or lower than white supervisors.

Nevertheless, supervisory evaluative data are considered to be of interest in the present context. Moreover, the black and the white "low aptitude" recruits have been equated on both the usual Navy tests and the miniature job learning tests. The followup criterion referenced performance data will also be available for both racial groups. Accordingly, the opportunity will exist for checking this response bias conjecture, found in industry, in the Navy context.

Criterion Tests

To develop the criterion tests for use in this study, Applied Psychological Services' personnel reviewed first portions of NAVPERS 18068A relevant to the machinist mate rate. Several proposed ideas for criterion referenced six month Fleet performance tests were extracted. The next step was to elaborate on these testing possibilities with experts in the machinist mate rate. The experts in this situation were primarily needed to supply scorable testing items, fruitful testing suggestions, and statements of criteria for "acceptable" work on each criterion objective after six months experience as a machinist mate striker. Four such experts were made available by the Naval Damage Control Training Center, Philadelphia.

---

* Three Chief Warrant Officers and One Master Chief Petty Officer.

These experts agreed that the following practical performance items would constitute an adequate test of the ability of a machinist mate striker, with six months Fleet experience, to meet normal performance expectations:

1. standing messenger watch
2. making and breaking a flange
3. packing a valve
4. procedures in common malfunction and in emergency situations
5. knowledge of use and names of common equipment and tools
6. general alertness and common sense in the work situations

Three separate meetings were held to isolate, derive, and define the performance objectives and the methods for measuring performance on these objectives. As a result, six nonverbal performance tests were defined. Each of these is described categorically below.

Message Watch

The message watch examination is a test of the examinee's ability to record data accurately from a throttle board and to determine malfunctions indicated by the data. Pictorial and simulated throttle board situations are presented, one at a time, to the examinee. The examinee is required to record accurately data from the throttle board pictures and to report to the examiner those throttle board readings which indicate a malfunction in the system. The examinee receives two scores. One score is based on his degree of accuracy in recording data from the simulated throttle board picture; the second score indicates his ability to detect malfunctions from the data. A sample item is presented as Figure 4-1.

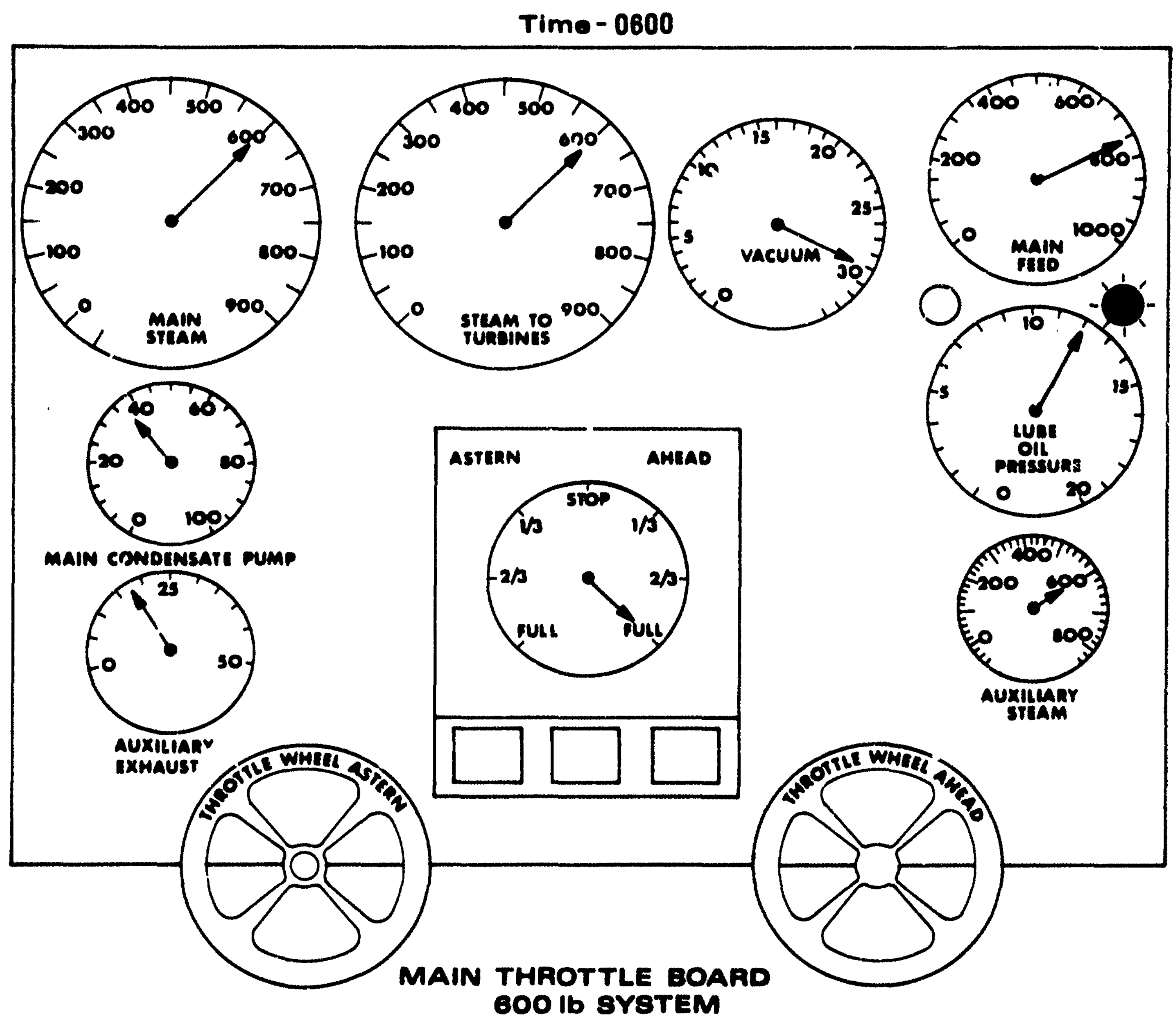

Figure 4-1. Sample throttle board test item.

Making-Breaking a Flange

The making-breaking a flange test is an individually administered performance test. The examinee is required to break and make a flange using the following tools and items: (a) a gasket, (b) one assembled six inch flange, (c) one scraper, (d) two box end wrenches, and (e) one rag. The flange on which the test is performed is shown in Figure 4-2. Scoring is based on following the correct procedures, care and use of tools, and adherence to safety regulations.

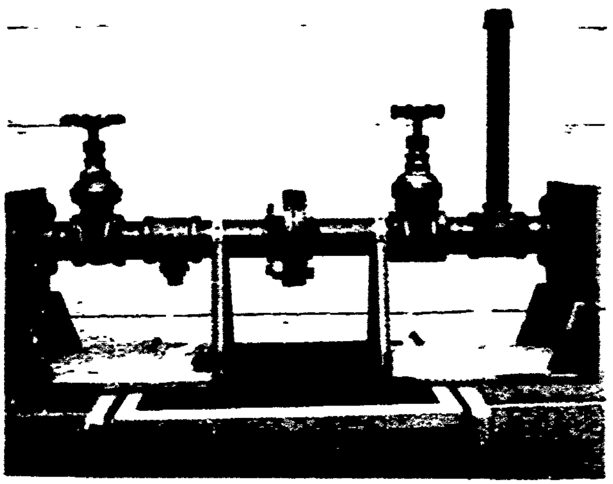

Figure 4-2. Flange for flange making-breaking test.

Packing a Valve

The valve packing test is also an individually administered performance test. The examinee is required to pack a valve using: (a) a large, mounted valve, (b) packing material, (c) a knife, (d) box end wrenches, (e) a packing puller, and (f) one very large adjustable wrench. Scoring is based on following the correct procedures, care and use of tools, and adherence to safety regulations.

Malfunction and Emergency Procedures

The malfunction and emergency procedures test is an individually administered test. Each item in the test consists of a set of pictures depicting a common emergency or malfunctive correction sequence. The task of the examinee is to place the pictures, which are presented in scrambled order, in the correct sequence. To do this the examinee must, of course, first recognize what is being represented. Each item in the set consists of a situation which is critical to or frequently encountered in machinist mate performance. The various problems depicted (in order of difficulty) in the picture arrangement test are: (a) electric shock, (b) fire in compartment, (c) leaky valve repair, (d) tank gauging, (e) fire hose assembly, (f) ruptured pipe, and (g) spring bearing repair.

Scoring is based on the number of pictures in each item set which are placed in the correct order. A time bonus is given for fast performance. The number of pictures in each set ranges from four to six. A sample item is presented in Figure 4-3.

Equipment/Tools Names and Use

The equipment/tools names and use test is also individually administered. The items consist of a set of cards. Each card shows a typical machinist mate work situation and three tools which might be employed to complete the job. The items are arranged in a hierarchical order of difficulty. For each item, the examinee is required to select, from among the three shown, the best tool for completing the specific job depicted. Extra credit is awarded if the examinee can also state verbally the correct name of the tool. A sample item is shown in Figure 4-4.

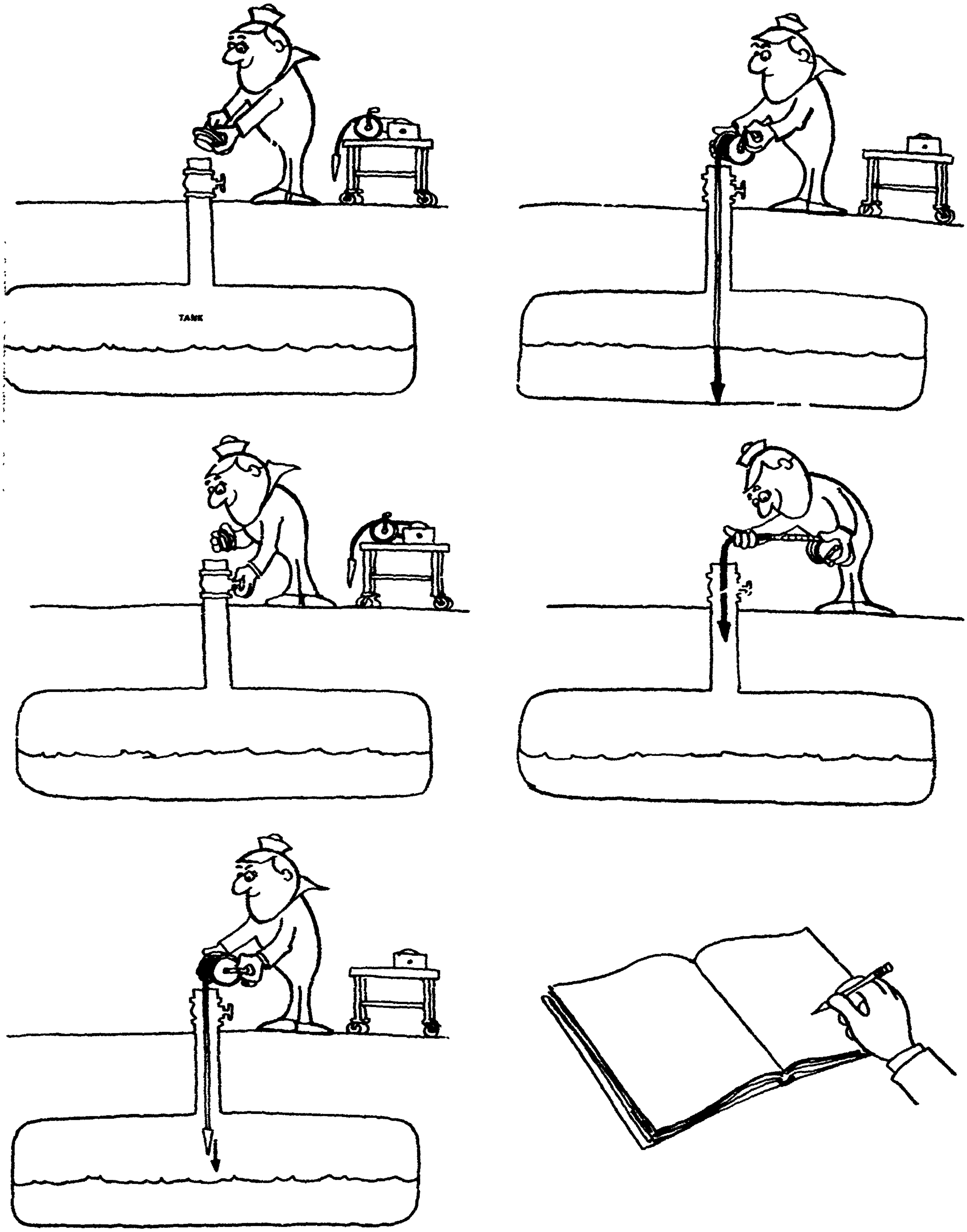

Figure 4-3. Sample malfunction and emergency procedure item (arranged in correct sequence).

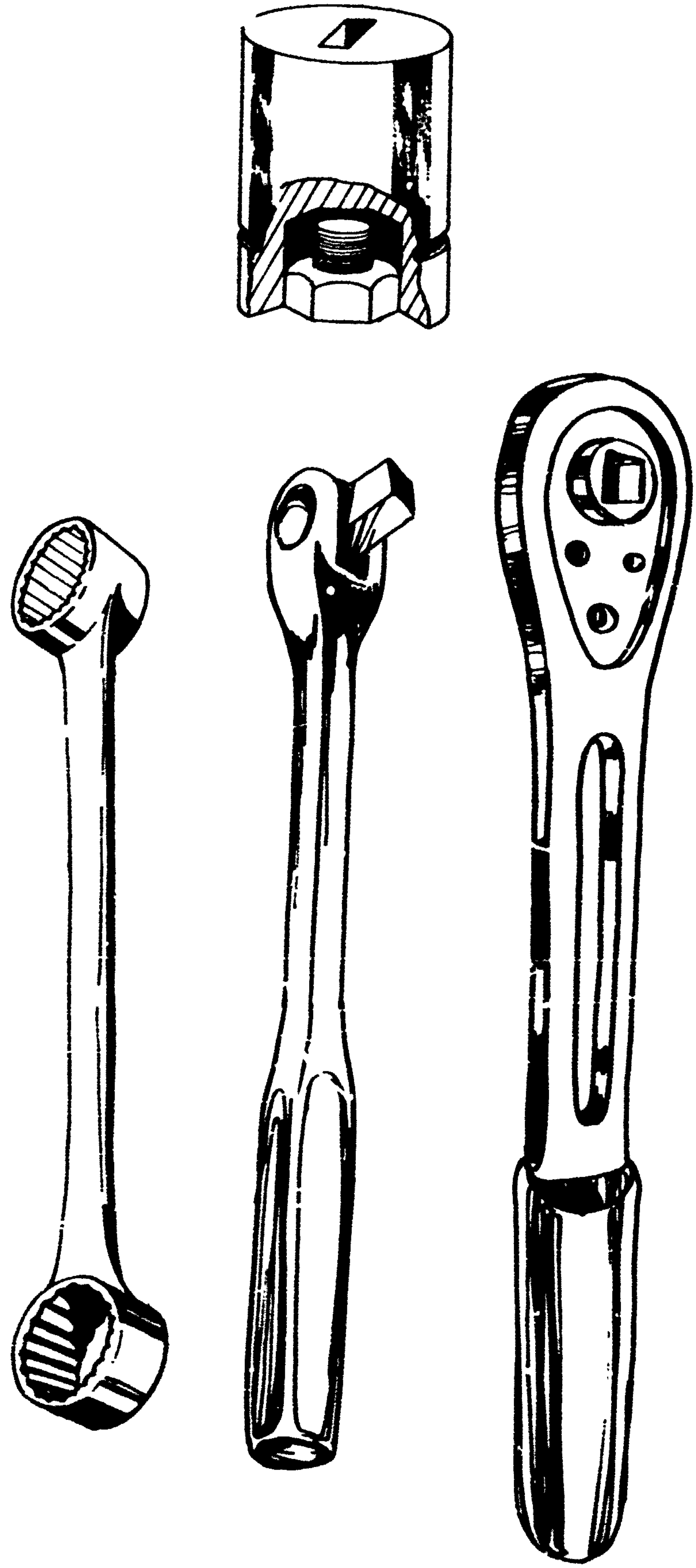

Figure 4-4. Sample equipment/tools names and use item.

General Alertness/Common Sense

The general alertness/common sense test is also a pictorial, individually administered test. Each item consists of a picture of a typical machinist mate work situation in which the sailor shown is doing something wrong. The task of the examinee is to detect and report what is wrong in each picture. The items are arranged in ascending order of difficulty and the scoring is based on the number of correct responses. A sample item is shown in Figure 4-5.

Validation of Miniature Job Sample Learning Tests

Once the criterion data are collected, Applied Psychological Services will determine the extent that the miniature job sample learning tests predict the criterion scores. The working hypothesis is that the miniature job learning tests, on a collective basis, will predict criterion performance better than the usual Naval classification tests for both the white and the black "low aptitude" recruits. We also hypothesize that no differential validity will be involved and that significantly different criterion test scores are anticipated for our "probably acceptable" and the "probably unacceptable" groups. As stated previously, supervisor evaluative data will be collected, and supervisory interviews will be conducted. These interview and rating data will provide further insight into the ability of those who have passed the miniature job learning tests but who scored below the cut point for the machinist mate rating on the usual Navy classification tests. Finally, the criterion tests will be administered to a sample of machinist mates who possess the same experience in the Fleet as our "low aptitude" sample but who have graduated from the machinist mate "A" school.

The end product will be a method for identifying those persons who can be successful in the Fleet who might otherwise be eliminated from consideration. The results may suggest a dualistic testing approach in which persons scoring below cut points on the usual Navy tests are given a "second chance" through a battery of nonverbal miniature job learning tests.

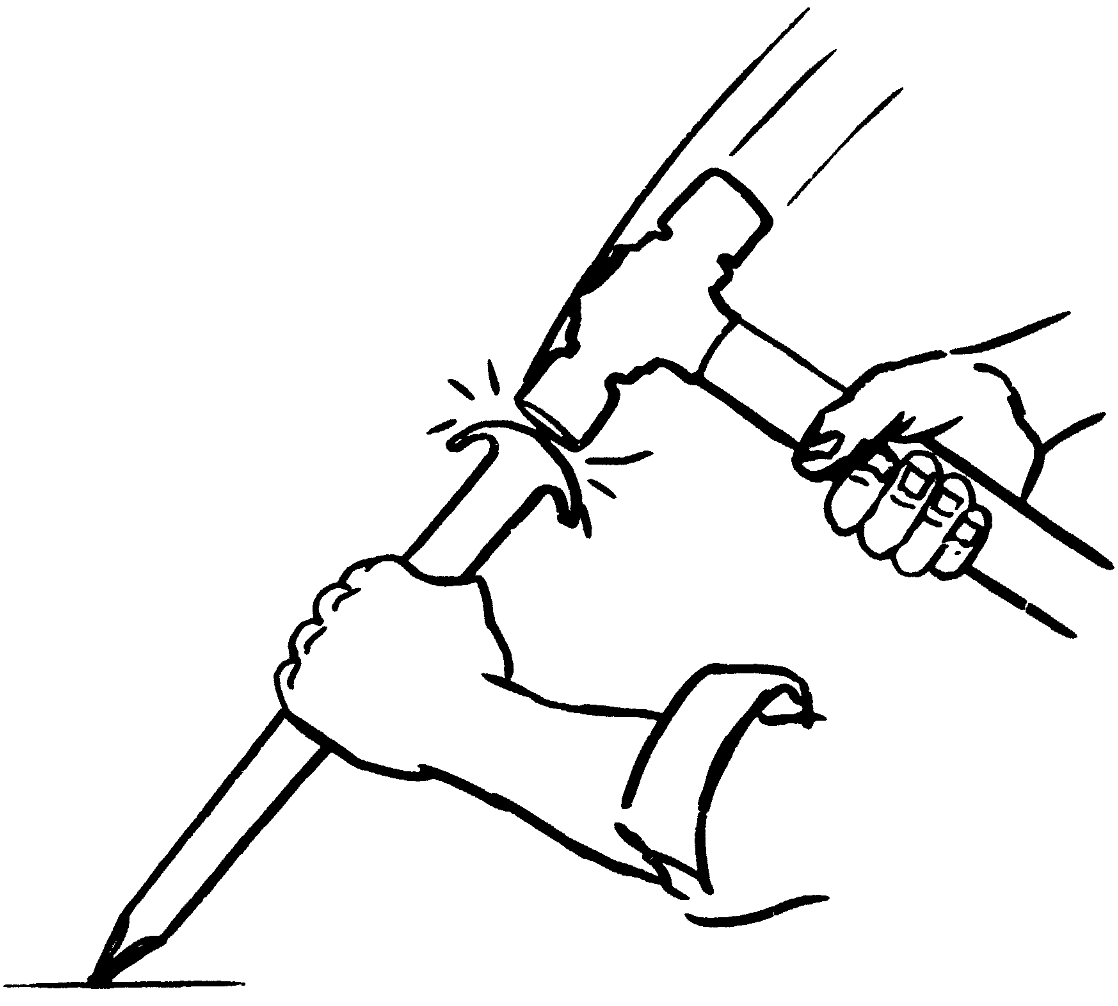

Figure 4-5. Sample general alertness/common sense item.

CHAPTER V

SUMMARY AND CONCLUSIONS

This study was designed to develop a nonverbal, culture fair assessment procedure which identifies Navy recruits who can learn to perform Navy jobs in skilled ratings even though they fail to meet the usual selection standards as measured by the current Navy classification tests. The underlying hypothesis was that recruits, judged by the usual classification methods to be of "low aptitude" but who exhibit the ability to learn and perform sample aspects of a Navy job, will be able to learn and perform on the total job--provided proper on-the-job training is given.

The machinist mate rating was used as a basis for testing the theoretic and methodological concepts involved. A sample of 50 white and 49 black "low aptitude" recruits took part in a series of miniature job learning and evaluative situations. These situations sampled those jobs which are frequently performed by machinist mates during their first six months in the Fleet or which are critical to performance during this period. At the conclusion of each job learning session, a skill based, nonverbal performance test was administered to each recruit. The recruits also: (1) completed a personal background questionnaire, and (2) participated in a post-test interview which inquired into reactions to the training and testing methods involved.

The personal background questionnaire was also administered to 118 persons who met the usual qualifying standards for the machinist mate rate. Those "low aptitude" recruits, to whom the miniature job learning instructional tests were administered, have been placed aboard ships of the Atlantic Fleet for entry into the machinist mate rating.

A battery of criterion referenced performance tests has been developed for measuring the level of Fleet achievement of those "low aptitude" persons who, on the basis of the miniature job sample learning test results, were classified as "probable acceptable" or "probable unacceptable" in terms of predicted Fleet performance.

The scores of the "low aptitude" sample on the miniature job learning tests, along with the scores of the sample on the usual Navy classification tests were intercorrelated. These data, along with interscorer reliability data and measures of central tendency, were used as the primary basis for preliminary evaluation of the miniature job learning tests.

Preceding page blank

The personal history questionnaire was factor analyzed, and the factor scores of both the "low aptitude" and the "A" school sample were compared. Additionally, the factor scores of the "low aptitude" sample were correlated with their scores on the mini learning tests and on the usual Navy classification tests. The results of the work, completed to date, suggest the following conclusions:

1. The six miniature job learning tests measure factors which are independent from the factors measured by the usual Naval classification tests.

2. The white group and the black group performed equally well on the miniature job learning tests; accordingly, the tests are considered to be reasonably "culture fair."

3. Adequate interscorer reliability was demonstrated.

4. Statistically significant differences were evidenced between the factor scores of the "low aptitude" sample and the factor scores of the "A" school sample on eight of the nine cultural deprivation factors extracted from the cultural deprivation questionnaire.

5. The mini job learning tests are less culturally loaded than the usual Navy classification tests.

6. The type of testing program here described is preferred by "low aptitude" personnel over the usual type of classification testing performed in the Navy.

REFERENCES

American Psychological Association Task Force on Employment Testing of Minority Groups. Job testing and the disadvantaged. American Psychologist, 1969, 24, 637-650.

Aronson, J. The use of a moderator variable to key a biographical inventory. Unpublished Doctoral Dissertation, New York University, 1967.

Atkinson, J. Motivational determinants of risk taking behavior. In J. Atkinson and N. Feather (Eds.), A theory of achievement motivation. New York: Wiley, 1966. Pp. 11-30.

Baehr, N., Saunders, D., Froemel, E., & Furcon, J. The prediction of performance for black and for white police patrolmen. Professional Psychology, 1971, 2(1), 46-57.

Bartlett, C., & O'Leary, B. A differential prediction model to moderate the effects of heterogeneous groups in personnel selection and classification. Personnel Psychology, 1969, 22(1), 1-18.

Battle, E., & Rotter, J. Children's feelings of personal control as related to social class and ethnic group. J. of Personality, 1963, 31, 482-490.

Bennett, G. Factors affecting the value of validation studies. Personnel Psychology, 1971 (In Press).

Boehm, V. Negro-white differences in validity of employment and training selection procedures: Summary of research evidence. J. of Applied Psychology, 1971 (In Press).

Burger, W. Griggs, et al., vs. Duke Power Company. Supreme Court of the United States, No. 124 - October Term, 1970 (March 8, 1971).

Cole, M. & Bruner, J. Cultural differences and inferences about psychological processes. American Psychologist, 1971, 867-876.

Crandall, V., Katkovsky, W., & Preston, A. Motivational and ability determinants of young children's intellectual achievement behaviors. Child Development, 1962, 33, 643-661.

Crandall, V., Preston, A., & Rabson, A. Maternal reactions and the development of independence and achievement behavior in young children. Child Development, 1960, 31, 243-251.

Darlington, R. Another look at 'cultural fairness". J. of Educational Measurement, 1971, 867-876.

DuBois, P. A general summary of psychological research in support of technical training. Office of Naval Research Contract No. Nonr-816(14) Naval Air Technical Training, 1970.

Dunnette, M. A modified model for test validation and selection research. J. of Applied Psychology, 1963, 47, 317-323.

Einhorn, H., & Bass, A. Methodological considerations relevant to discrimination in employment testing. Psychological Bulletin, 1971, 75(4), 261-269.

Enneis, W. Minority employment barriers from the EEOC viewpoint. Paper presented at the annual meeting of the American Psychological Association, September, 1969.

Enneis, W. Personnel testing and equal employment opportunity. Paper presented at the annual meeting of the Pennsylvania Psychological Association, June, 1971.

Farr, J., O'Leary, B., Pfeiffer, C., Goldstien, I. & Bartlett, C. Ethnic group membership as a moderator in the prediction of job performance: An examination of some less traditional predictors. Pittsburgh, Pa: American Institutes for Research, 1971.

Flaugher, R., Campbell, J., & Pike, L. Prediction of job performance for negro and white medical technicians. Project Report 69-5. Princeton, N.J.: Educational Testing Service, 1969.

Foley, F. Validity of the OQT for minority group applicants to officer candidate school. Naval Personnel Research and Development Laboratory. January 1971, WRR-71-1 (Revised).

Fox, W., Taylor, J., & Caylor, J. Aptitude level and the acquisition of skills and knowledges in a variety of military training tasks. HumRRO Technical Report 69-6, May 1969.

Friedrichs, A., Hertz, T., Moynahan, E., Simpson, W., Arnold, N., Christy, N., Cooper, C., & Stevenson, H. Interrelations among learning and performance tasks at the preshcool level. Developmental Psychology, 1971, 4, 173-177.

Guinn, N., Tupes, E., & Alley, W. Cultural subgroup differences in the relationships between Air Force aptitude composites and training criteria. AFHRL-TR-70-35. Lackland AFB, Tex. Personnel Research Division, Air Force Human Resources Laboratory, September, 1970.

Guion, R. Personnel selection. Annual Review of Psychology, 1967, 18, 191-216.

Grant, D., & Bray, D. Validation of employment tests for telephone company installation and repair occupations. J. of Applied Psychology, 1970, 54(1), 7-14.

Grunzke, N., Guinn, N., & Stauffer, G. Comparative performance of low ability airmen. AFHRL-TR-70-4. Texas: Lackland Air Force Base, 1970.

Hess, R., & Shipman V. Early experience and the socialization of cognitive modes in children. Child Development, 1965, 36, 869-886.

Hooprich, E. A second investigation of the feasibility of Navy commissary-man training for group IV personnel. San Diego, Calif.: Naval Personnel Research Activity, 1968.

Jensen, A. How much can we boost IQ and scholastic achievement? Harvard Educational Review, 1969, 39(1), 1-123.

Katkovsky, W., Crandall, V., & Good, S. Parental antecedents of children's beliefs in internal-external control of reinforcements in intellectual achievement situations. Child Development, 1967, 38, 765-776.

Katz, I. Some motivational determinants of racial differences in intellectual achievement. International Journal of Psychology, 1967, 2(1), 1-12.

Kimble, S. A measure of cultural deprivation. Dissertation Abstracts, 1971, 31, (91A), 4552.

Kirkpatrick, J., Ewen, R., Barrett, R., & Katzell, R. Testing and fair employment Fairness and validity of personnel tests for different ethnic groups. New York: NYU Press, 1968.

Krug, R. Suggested approaches for test development and measurement. Personnel Psychology, 1966, 19(1), 24-35.

Linn, R., & Werts, C. Considerations for studies of test bias. J. of Educational Measurement, 1971, 8(1), 1-4.

Lopez, F. Current problems in test performance of job applicants: 1 Personnel Psychology, 1966, 19(1), 10-18.

McFann, H. Progress report on HumRRO research on project 100,000. HumRRO Professional Paper, 1969, No. 25-69.

Mingione, A. Need for achievement in negro and white children. J. of Consulting Psychology, 1965, 29, 108-111.

Mitchell, N., & Albright, L. Biracial validation of selection procedures in a large southern plant. Proceedings of the 76th Annual Convention of the APA, 1968, 3, 575-576.

Moore, C., MacNaughton, J., & Osburn, H. Ethnic differences within an industrial selection battery, Personnel Psychology, 1969, 22, 473-482.

Morgan, J. The achievement motive and economic behavior. In J.W. Atkinson and N.T. Feather (Eds.), A theory of achievement motivation. Wiley New York, 1966, Pp. 11-30.

Plag, J., & Goffman, J. The Armed Forces Qualification Test Its validity in predicting military effectiveness for Naval enlistees. Personnel Psychology, 1967, 20(3), 328-340.

Rimland, B. The search for measures of practical intelligence Research on project 100,000. Paper presented at the annual meeting of the American Psychological Association, September, 1969.

Rohwer, W., Ammon N., Suzuki, N., & Levin, J. Population differences and learning proficiency. J. of Educational Psychology, 1971, 62(1), 1-14.

Rosen, B. The achievement syndrome: A psychocultural dimension of social stratification American Sociological Review, 1956, 21, 203-211.

Rosen, B. Race, ethnicity, and the achievement syndrome. American Sociological Review, 1959, 24, 47-60.

Rosenhan, D. Effects of social class and race on responsiveness to approval and disapproval. J. of Personnel and Social Psychology, 1966, 4, 253-259.

Rotter, J. Generalized expectancies for internal vs. external control of reinforcement. Psychological Monograph, 1966, 80, No. 1 (whole No. 609).

Ruch, F. What impact did the Supreme Court decision in the D Case have on employment procedures? An address delivered to the Edison Electric Institute, Industrial Relations Committee at the Jack Tarr Hotel, San Francisco, May, 1971.

Ruda, E., & Albright, L. Racial differences on selection instruments related to subsequent job performance. Personnel Psychology, 1968, 21, 31-41.

Rundquist, E. The prediction ceiling. Personnel Psychology, 1969, 22, 109-116.

Scott, J., & Phelan, J. Expectancies of unemployable mates regarding source of control of reinforcement. Psychological Reports, 1969, 25(3), 911-913.

Seeman, N. Alienation and social learning in a reformitory. American Journal of Sociology. 1963, 69, 270-284.

Seeman, N., & Evans, J. Alienation and learning in a hospital activity. American Sociological Review, 1962, 27, 772-782.

Semler, I., & Iscoe, I. Comparative and developmental study of the learning abilities of negro and white children under poor conditions. J. of Educational Psychology, 1963, 54(1), 38-44.

Shrivasta, P., & Tiwari, N. Socioeconomic stratification of real achievement. Psychological Studies, 1967, 12(1), 9-16.

Slocum, J., & Strawser, R. Racial differences in job attitudes. Experimental Publishing System. June, 1971, 12, NS. No. 471-12.

Standlee, L., & Saylor, J. Second study of equipment operator class "A" school training for group IV personnel. San Diego, Calif.: Naval Personnel Research Activity, 1969.

Sticht, T. Learning by listening in relation to aptitude reading, and rate controlled speech. HumRRO Technical Report 69-23, 1969.

Taylor, J., Montague, E., & Hauke, R. The interrelationships of ability level, instructional system, and skill acquisition. HumRRO Professional Paper, 1970, No. 29-70.

Thorndike, R. Concepts of culture fairness. J. of Educational Measurement, 1971, 8(2), 63-70.

Van Matre, N. An instrument reading training course for group IV personnel. San Diego, Calif.: Naval Personnel and Training Research Laboratory, 1971. Report No. SRR-71-26.

Van Matre, N., & Steineman, J. A performance oriented electronics technician training program. II. Initial Fleet followup evaluation of graduates. San Diego, Calif.: U.S. Naval Research Laboratory, Technical Bulletin STB 67-157, 1966.

Van Matre, N., & Harrigan, R. A performance oriented electronic technician training program. V. Final Fleet followup of graduates. San Diego, Calif.: U.S. Naval Training Research Laboratory, Research Report SRR 70-15.

Vineberg, R., Sticht, T., Taylor, E., & Caylor J. Effects of aptitude (AFQT), job experience, and literacy on job performance: Summary of HumRRO workunits UTILITY and REALISTIC. HumRRO, August, 1970.

Wollowick, H., Greenwood, J., & McNamara, W. Psychological testing with a minority group population. Proceedings of the 77th Annual Convention of the American Psychological Association, 1969, 4, 609-610.

APPENDIX A

LESSON IV

Training Evaluation
Equipment Operation
Examiner Instructions

Task - This is a test of the examinee's ability to start up and shut down a motor.

Test Materials - The following materials should be provided:

1. Two electric motors
2. Two long dipsticks
3. Two short dipsticks
4. Two rags
5. Two oil cans filled with red water
6. Two fuel cans filled with green water
7. Two battery boxes with extension cords
8. Scoring checklist for each student. Make sure the student's name is on the checklist before you begin testing him.
9. Two stopwatches

General Directions to Examiner:

1. Each examiner will be tested individually in a quiet, well lighted room or area.
2. All students, other than the examinee, must be kept outside of the test room while testing is in progress.
3. Read very carefully and thoroughly the "Examinee Instructions." Be certain you understand the test and the method for administering it before attempting to give it to the examinee.
4. Make certain each examinee understands his instructions before he begins the test. Supply no information beyond what is needed for understanding the test procedure. Remember, this is a test situation, not a training one.
5. Make certain that all examinees who have eyeglasses are wearing them.
6. Allow each subject 10 minutes to start up and shut down the motor.
7. After you have finished testing one subject, he is to remain quietly in the testing room out of view (behind) the next testee.

Directions to Examinees (to be read on tape)

"THIS IS A TEST OF YOUR ABILITY TO START UP AND SHUT DOWN A MOTOR. MACHINIST MATES PERFORM TASKS VERY MUCH LIKE THIS."

"I WANT YOU TO START UP AND SHUT DOWN THIS MOTOR IN EXACTLY THE SAME WAY AS YOU WERE TAUGHT. TRY TO PERFORM EVERY STEP EXACTLY AS WE SHOWED YOU DURING THE LESSON. YOU WILL ONLY RECEIVE FULL CREDIT IF YOU PERFORM THE STEPS CORRECTLY AND IN THE RIGHT ORDER. IF YOU DO NOT PERFORM THE STEPS IN THE RIGHT ORDER, YOU WILL NOT RECEIVE FULL CREDIT."

"YOU WILL HAVE TEN MINUTES TO COMPLETE THIS TASK."

"BEGIN."

Scoring of Equipment Operation Checklist

1. Place a yes after each item that is performed correctly.
2. Place a no after each item that is performed incorrectly.
3. Leave the space blank if the step is not performed.
4. Encircle each correctly performed but out of sequence step.
5. Allow two points for each correctly performed step.
6. Allow one point for each correctly performed (circled) step out of sequence, except for reversals of screwing on the oil and fuel tank lids.
7. Do not allow any points for incorrect steps or for steps not performed.

LESSON IV

Training Procedure for Equipment Operation

To the Instructor: Keep this sheet in front of you at all times during this training session. Do and say exactly what is said on these pages.

Procedure

1. Stand next to the motor.

2. Say: "I HAVE IN FRONT OF ME A SMALL MOTOR WITH A PUMP ATTACHED TO IT. THERE ARE SEVERAL STEPS I MUST PERFORM IN ORDER TO START UP THIS MOTOR."

3. Say: "MY FIRST STEP IS TO REMOVE ANY OILY RAGS OR TRASH LYING ON OR NEAR THE MOTOR. YOU CAN ALL SEE THAT THERE IS A RAG NEXT TO THE MOTOR. I WILL REMOVE IT. IF I DIDN'T REMOVE THE RAG, THERE IS A CHANCE THAT A SPARK FROM THE MOTOR WOULD IGNITE THE RAG AND CAUSE A FIRE."

4. Remove rag.

5. Say: "THE SECOND STEP IS TO CHECK THE FUEL LEVEL USING A LONG DIPSTICK. REMEMBER TO USE THE LONG DIPSTICK WHEN YOU CHECK THE FUEL LEVEL."

6. Hold up long dipstick.

7. Say: "THIS IS THE FUEL TANK."

8. Point to fuel tank.

9. Say: "I WILL UNSCREW THE FUEL TANK LID, CLEAN OFF THE DIPSTICK, AND PUT THE DIPSTICK IN UNTIL IT TOUCHES THE BOTTOM OF THE FUEL TANK."

10. Unscrew top of fuel tank, clean the dipstick, and put long dipstick in until it touches bottom. Pull dipstick out of fuel tank and hold it up to the class.

11. Say: "NOTICE THAT THE FUEL LEVEL IN THE FUEL TANK IS NOT FULL. THE FUEL LEVEL MUST REACH THIS WHITE LINE TO BE FULL."

12. Point to white line.

13. Say: "SINCE THE FUEL TANK IS NOT FULL, I WILL ADD FUEL FROM THE FUEL STORAGE TANK WITH THIS FUEL FUNNEL."

14. Hold up fuel funnel to class. Point to fuel storage tank. Insert funnel in fuel tank. Unscrew fuel storage tank. Pour fuel from fuel storage tank into fuel tank. Put fuel storage tank aside. Put lid on fuel storage tank.

15. Say: "REMEMBER THAT AFTER YOU POUR FUEL YOU MUST SCREW THE LID BACK ONTO THE FUEL STORAGE TANK. NOW I WILL AGAIN CHECK THE FUEL LEVEL WITH THE DIPSTICK. BEFORE I CHECK THE FUEL LEVEL, I MUST WIPE OFF THE DIPSTICK WITH A RAG."

16. Check fuel level with long dipstick, after wiping it off with a rag.

17. Say: "YOU CAN NOW SEE THAT THE FUEL LEVEL HAS REACHED THE WHITE LINE. MY NEXT STEP IS TO PUT THE LID BACK ONTO THE FUEL TANK."

18. Put lid back on fuel tanks.

19. Say: "NOW I MUST SEE IF THE OIL LEVEL IS CORRECT. I WILL CHECK THE OIL LEVEL IN THE SAME WAY AS I CHECKED THE FUEL LEVEL NOTICE THAT I USE A SHORT DIPSTICK RATHER THAN A LONG DIPSTICK, AN OIL STORAGE CAN RATHER THAN A FUEL STORAGE CAN, AND AN OIL FUNNEL RATHER THAN A FUEL FUNNEL."

20. Point to these three items and mention their names again: "SHORT DIPSTICK, OIL STORAGE CAN, AND OIL FUNNEL."

21. Say: "THE OIL IS NEEDED TO LUBRICATE THE MOTOR AND THE PUMP."

22. Say: "NOW I WILL CHECK THE OIL LEVEL USING THE SHORT DIPSTICK."

23. Hold up short dipstick.

24. Say: "THIS IS THE OIL TANK."

25. Point to oil tank.

26. Say: "I UNSCREW THE OIL TANK LID, WIPE OFF THE DIPSTICK, AND PUT THE DIPSTICK IN UNTIL IT TOUCHES THE BOTTOM OF THE OIL TANK."

27. Unscrew top of oil tank and put short dipstick in until it touches bottom. Pull dipstick out of oil tank and hold it up to the class.

28. Say: "NOTICE THAT THE OIL LEVEL IN THE OIL TANK IS LOW. THE OIL LEVEL MUST REACH THIS WHITE LINE."

29. POINT TO WHITE LINE.

30. Say: "SINCE THE OIL TANK IS NOT FULL, I WILL ADD OIL FROM THE OIL STORAGE TANK WITH THIS OIL FUNNEL."

31. Hold up oil funnel to class. Point to oil storage tank. Insert funnel in oil tank. Unscrew oil storage tank. Pour oil from oil storage tank into oil tank. Put oil storage tank aside. Put lid on oil storage tank.

32. Say: "REMEMBER, AFTER YOU POUR OIL YOU MUST SCREW THE LID BACK ONTO THE OIL STORAGE TANK. NOW, I AGAIN CHECK THE OIL LEVEL WITH THE DIPSTICK. BEFORE CHECKING THE OIL LEVEL, I AGAIN WIPE THE DIPSTICK OFF WITH A RAG."

33. Check oil level with short dipstick.

34. Say: "YOU CAN NOW SEE THAT THE OIL LEVEL HAS REACHED THE WHITE LINE. MY NEXT STEP IS TO PUT THE LID BACK ONTO THE OIL TANK."

35. Put lid back on oil tank.

36. Say: "NEXT I WILL CHECK TO SEE IF THIS VALVE IS IN THE BYPASS POSITION."

37. Point to valve.

38. Say: "THE VALVE IS NOW IN THE BYPASS POSITION, IF NO WATER IS FLOWING INTO THE BUCKET."

39. Point to inside of bucket.

40. Say: "NEXT, WE WILL CONNECT THE BATTERY WITH THIS PLUG. THIS IS THE BATTERY AND THIS IS THE PLUG."

41. Point to battery and plug. Put plug into battery socket.

42. Say: "NOTICE THAT I PUT THE PLUG INTO THE BATTERY JUST AS IF I WERE PLUGGING IN A LAMP."

43. Say: "NOW I WILL TURN ON THE IGNITION SWITCH. THIS IS THE IGNITION SWITCH."

44. Point to ignition switch and turn it on.

45. Say: "MY FINAL STEP IS TO TURN THE BYPASS VALVE ONTO THE PUMP."

46. Turn bypass valve so that water pumps.

47. Say: "NOW WE MUST LEARN HOW TO SHUT THIS MOTOR OFF. ALL WE HAVE TO DO IS TURN THE VALVE BACK TO THE BYPASS POSITION. NO WATER SHOULD BE GOING INTO THE BUCKET."

48. Turn valve to bypass.

49. Say: "AND TURN THE IGNITION OFF."

50. Turn off ignition.

51. Say: "YOUR FINAL STEP IS TO WIPE OFF THE FUNNELS WITH A RAG AND WIPE UP ANY OIL AND FUEL SPILLS THAT OCCURRED."

52. Say: "WE WILL GO OVER THIS ONCE AGAIN BEFORE WE ALLOW YOU TO PRACTICE WHAT YOU HAVE LEARNED.

53. Empty fluid from oil and fuel tanks back into their storage containers. Perform steps 1-52 again.

54. Allow each man 10 minutes to practice starting up (the other students can watch) and shutting off motor. Observe each man closely. Help them or prompt them whenever they are having difficulty. Correct mistakes.

LESSON IV

Scoring Checklist for Equipment Operation

Name	Date

1. Removes rag from under motor ________
2. Unscrews fuel tank lid ________
3. Wipes dipstick clean ________
4. Inserts long dipstick ________
5. Observes need for fuel ________
6. Does not wipe dipstick ________
7. Adds fuel from fuel storage tank ________
8. with fuel funnel ________
9. Cleans long dipstick ________
10. Inserts long dipstick again ________
11. Observes adequate fuel level ________
12. Does not wipe dipstick ________
13. Puts lid back on fuel storage tank ________
14. Puts lid back on fuel tank ________
15. Unscrews oil tank lid ________
16. Wipes short dipstick clean ________
17. Inserts short dipstick ________
18. Does not wipe dipstick ________

19. Adds oil from oil storage tank ______
20. with oil funnel ______
21. Cleans short dipstick ______
22. Inserts short dipstick again ______
23. Does not wipe dipstick ______
24. Puts lid back on oil storage tank ______
25. Puts lid back on oil tank ______
26. Checks to see if valve is in bypass position spot (by looking in bucket) ______
27. Puts plug in battery socket ______
28. Turns on ignition switch ______
29. Turns bypass valve on (Prompt him to shut it off if he doesn't do it immediately) ______
30. Turn valve to bypass ______
31. Turn off ignition ______
32. Does not pull plug ______
33. Cleans up funnels and spills of oil and fuel ______

Total Points ______

LESSON III

Assembly

Scoring Checklist

______________________ ______________

Name Date

1. Takes packing nut and stem ______
2. Screws packing nut to top of stem ______
3. Takes gate ______
4. Screws gate on bottom of stem ______
5. Winds gate all the way up stem ______
   (Prompt if Step 5 performed incorrectly)
6. Inserts gate and stem assembly into body of valve ______
7. Screws ate and stem assembly on to body of valve ______
8. Inserts handle onto top of stem ______
9. Screws handle onto top of stem with handle nut ______
10. Screws on first 3/4" nipple ______
11. Screws on second 3/4" nipple ______
12. Checks assembled valve to see if parts are fitted tightly ______

Total plus ______

Total minus ______

Total ______

LESSON II

Name of Trainee ____________________ Date ____________

Scoring Checklist for Making a Full Face Gasket

1. Uses round end ______
2. of ball peen hammer ______
3. and correct size gasket material ______
4. and taps out one bolt hole ______
5. lightly. ______
6. Inserts bolt in hole. ______
7. Taps out diagonally opposite bolt hole and inserts bolt. ______
8. Taps out remaining bolt holes in any order ______
9. lightly. ______
10. Taps out inside circumference of flange ______
11. using round end of ball peen hammer ______
12. lightly. ______
13. Taps at outside circumference of flange ______
14. using flat end of ball peen hammer ______
15. lightly. ______
16. Remove tapped out pieces of gasket material from gasket and flange. ______
17. No frayed edges in final product. ______
18. No gouges or scratches in final product. ______
19. Does not ruin or throw away any pieces of gasket material. ______